



UNIVERSIDAD CARLOS III DE MADRID
ESCUELA POLITÉCNICA SUPERIOR

DEPARTAMENTO DE INFORMÁTICA
DOCTORADO EN CIENCIA Y TECNOLOGÍA INFORMÁTICA

TESIS DOCTORAL

Modelado y Desarrollo en el Reconocimiento de Actividades de Grupo

Alberto Pozo Esteban

DIRIGIDA POR

Miguel Á. Patricio Guisado

Jesús García Herrero

Julio 2.016

This work is distributed under the Creative Commons 3.0 license. You are free to copy, distribute and transmit the work under the following conditions: (i) you must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work); (ii) you may not use this work for commercial purposes, and; (iii) you may not alter, transform, or build upon this work. Any of the above conditions can be waived if you get permission from the copyright holder. See <http://creativecommons.org/licenses/by-nc-nd/3.0/> for further details.



Address:

Grupo de Inteligencia Artificial Aplicada
Departamento de Informática
Universidad Carlos III de Madrid
Av. de la Universidad Carlos III, 22
Colmenarejo 28270 — Spain

Modelado y Desarrollo en el Reconocimiento de Actividades de Grupo

Autor: Alberto Pozo Esteban

Directores: Miguel Á. Patricio Guisado
Jesús García Herrero

Firma del Tribunal Calificador:

Nombre y Apellidos

Firma

Presidente: D.

Vocal: D.

Secretario: D.

Calificación:

Leganes, de de 2.016.

Mama, Papa, Alex.

Índice general

Abstract	vii
Resumen	ix
Agradecimientos	xi
1 Introducción	1
1.1 Organización del documento	2
I Estado del Arte	5
2 Reconocimiento de actividades	9
3 Conjuntos de datos	17
4 Selección de características	21
II Propuesta	25
5 Modelado y Diseño del Reconocimiento de Actividades	29
5.1 Descripción General del Problema	30
5.2 Modelado del Dominio 2 vs 2 en Baloncesto	32
6 Sistema de Reconocimiento de Actividades de Grupo	41
6.1 Explosión de Características	43
6.2 Selección de Características	45

6.2.1	Wrapper global	45
6.2.2	Wrapper binario	47
6.2.3	Filtro mRMR global	47
6.2.4	Filtro mRMR binario	48
6.3	Reconocimiento de la Actividad de Grupo	49
7	Inef12Basketball Dataset	55
7.1	Contexto	56
7.2	Datos	61
7.3	Funciones Matlab	64
III	Resultados Obtenidos	71
8	Sistema de Reconocimiento de Actividades de Grupo	75
8.1	Resultados Obtenidos	76
8.1.1	Resultados obtenidos aplicando Wrapper global	77
8.1.2	Resultados obtenidos aplicando Wrapper binario	78
8.1.3	Resultados obtenidos aplicando filtro mRMR global	79
8.1.4	Resultados obtenidos aplicando filtro mRMR binario	80
8.1.5	Resumen de los resultados	81
9	Conclusiones	83
A	Trabajos publicados	85
B	Protocolos y Métodos de Evaluación	87
B.1	Validación cruzada dejando uno fuera	87
B.2	Métricas	87
	Bibliografía	89

Índice de figuras

2.1	Clasificación del estado del arte en el reconocimiento de actividades	10
5.1	Coordenadas X e Y del equipo atacante	35
5.2	Coordenadas X e Y del resto de jugadores respecto del que lleva el balón	36
5.3	Coordenadas X e Y de la velocidad de un jugador	37
5.4	Componentes laplacianos del grafo formado por los cuatro jugadores	38
6.1	Flujo principal del entrenamiento del sistema	42
6.2	Flujo principal de la clasificación	42
6.3	Selección de características mediante wrapper global	46
6.4	Selección de características mediante wrapper binario	46
6.5	Entrenamiento del HMM global	50
6.6	Clasificación con el HMM global	51
6.7	Entrenamiento del HMM binario	53
6.8	Clasificación con el HMM binario	54
7.1	Pasar por delante	57
7.2	Pasar de tercero	57
7.3	Push	57
7.4	Flash lateral	58
7.5	Flash corto	59
7.6	Flash largo	59
7.7	Cambio	60

7.8	Dos contra uno	60
7.9	Plotter con la trayectoria de los jugadores (INEF12BasketballShowTrajectories())	66
7.10	Ejemplo de ejecución de la función ShowFrames	68
8.1	Matriz de clasificación para la primera iteración del wrapper global	77
8.2	Resumen de los resultados obtenidos con los diferentes algoritmos	81

Índice de cuadros

5.1	Lista de Características	34
7.1	Positions.txt	61
7.2	PlayerList.txt	62
7.3	Equivalencia entre identificador y papel	62
7.4	HasBall.txt	63
7.5	TeamActivity.txt	63
7.6	TeamActivityDictionary.txt	64
7.7	Equivalencia entre indetificador de jugada y explicación	64
8.1	Selección de características basadas en wrapper global	78
8.2	Selección de características basadas en wrapper binario	79
8.3	Precisión obtenida en cada iteración con wrapper binario	79
8.4	Selección de características basadas en filtro mRMR global	80
8.5	Selección de características basadas en filtro binario	80
8.6	Precisión obtenida en cada iteración con filtro binario	81

Abstract

GROUP activities recognition is a research branch at a time of a great boom, both academic and commercial levels. There are a lot of domains in which great utility problems can be addressed from this new perspective, such as team sports, defense systems, ethology, etc. In each and every of these domains we find a set of elements that are coordinated to carry out a combined action to chase a common goal.

First of all, it has been developed a comprehensive introduction to the current state of art . Given the large number of new approaches, it has been tried to emphasized the most influential investigations of this knowledge branch, exposing their strengths and possible weaknesses.

However, despite the great academic interest and huge trade boom recognition of group activities, there are a very limited number of datasets suitable for the development of new approaches, so academically it is difficult to develop new systems and perform a correct comparison with the other approaches.

Therefore, one of the contributions of this thesis consist of a new dataset specially developed for the problem of recognition of group activities, focused on the basketball domain. In it, all the requirements to develop a proper new recognition of group activities system and the subsequent testing are provided.

On the other hand, due to the main nature of the problem, recognition of group activities is a domain characterized by having a high dimensionality, this implies that systems sometimes have to work with a too wide number of variables.

Therefore, another contribution of this thesis involves the design, implementation and testing of a new recognition of group activities system, focused on the explosion and subsequent selection of features. This system will be tested in four different versions (with different feature selection strategies), achieving results versus the data set described before.

While the formalization of the system is independent of the subsequent classification performed, in this thesis it will be given an overview of a particular classification to test both

the formalization developed and the explosion and subsequent selection of features, thus being able to conclude the validity or not of the proposed formalization.

Resumen

EL reconocimiento de actividades de grupo es una rama de investigación que se encuentra en un momento de gran auge, tanto a nivel académico como comercial. Existen una gran cantidad de dominios en los que surgen problemas de gran utilidad que pueden ser afrontados desde esta nueva perspectiva, entre los que se encuentran algunos tales como los deportes de equipo, los sistemas de defensa, la etología, etc. En todos y cada uno de estos dominios nos encontramos con un conjunto de elementos que se coordinan para llevar a cabo una acción conjunta, persiguiendo así un objetivo común.

En primer lugar, se ha desarrollado una completa introducción al estado del arte actual, dado el gran número de nuevas aproximaciones se ha tratado de hacer especial hincapié en las investigaciones más influyentes de esta rama del conocimiento, exponiendo sus puntos fuertes así como sus posibles debilidades.

Sin embargo, pese al gran interés académico y el enorme auge comercial del reconocimiento de actividades de grupo, existen un número muy limitado de conjuntos de datos apropiados para el desarrollo de nuevas aproximaciones, por lo que a nivel académico resulta difícil desarrollar nuevos sistemas y realizar una correcta comparación con el resto de aproximaciones.

Por todo ello, una de las contribuciones de esta tesis doctoral consiste en un nuevo conjunto de datos especialmente desarrollado para el problema del reconocimiento de actividades de grupo, enfocado en el dominio del baloncesto. En él, se aporta todo lo necesario para el correcto desarrollo y posterior testeo de nuevos sistemas de reconocimiento de actividades de grupo. Por otro lado, debido a la propia naturaleza del problema, el reconocimiento de actividades de grupo es un dominio que se caracteriza por tener una alta dimensionalidad, esto implica que en ocasiones los sistemas tengan que trabajar con un número de variables demasiado amplio.

Por ello, otra aportación de esta tesis doctoral supone el diseño, implementación y prueba de un nuevo sistema de reconocimiento de actividades de grupo, centrado en la explosión y posterior selección de características. Dicho sistema será probado en cuatro versiones diferentes (con diferentes estrategias de selección de características). Obteniendo resultados frente al

conjunto de datos descrito anteriormente.

Si bien la formalización del sistema es independiente de la posterior clasificación que se realice, en este trabajo se dará una visión de conjunto con una determinada clasificación para poder poner a prueba tanto la formalización desarrollada, como la explosión y selección de características posterior. Pudiendo así concluir la validez o no de la formalización propuesta.

Agradecimientos

EN primer lugar quiero agradecer a mis tutores Miguel Á. Patricio y Jesús García tanto su esfuerzo y apoyo para realizar este trabajo, como que se fijaran en mí para emprender esta aventura académica en el GIAA. No puedo olvidarme tampoco de mi tutor durante el PFC y el PFM, Antonio Berlanga, así como del resto del departamento que siempre me hizo sentir como en casa.

En él tuve la suerte de coincidir con compañeros que ahora son mucho más, como: Gonzalo, Álvaro, Kike, Mirren, Rodri, Nayat, Javier, Luis, José... además de algún que otro "adoptado" como Ramón.

Siguiendo la línea de la universidad, pero esta vez como alumno, me vienen a la cabeza cientos de recuerdos increíbles que engloban a decenas de personas, no puedo nombrarlas a todas, pero tampoco puedo dejar de acordarme de algunos como: Acebes, Peña, Fede, Belén, Laura, Julen, Quike, Gino y tantos otros con los que he compartido horas de "estudio", mus, cafetería y "bolas de nieve". De esos años me llevo los mejores recuerdos, y la mejor pareja con la que uno puede soñar, de nuevo, gracias Laura.

Al remontarme un poco más me viene a la cabeza la gente del instituto, algunos de los mejores y de los que más orgulloso me siento por seguir manteniendo el contacto son: Kandela, Lucía, otra vez Julen, Muri, Fer-Antonio, José Miguel, Borja, etc. . .

Y por fin llegamos a "la Raya", amigos tan antiguos que han estado ahí siempre. Nunca seré plenamente consciente de la suerte que tuve de caer por allí, no sería en absoluto quien soy si no fuera por vosotros: Moreno, Tamara , Arturo, Silvia, Joseko, Gema, Rubén, Ari, José Luis, Tamara (La Rubia), Mónica, Sebas, Puli, Mario, Pablo, Noe, María 19, AlbertoP, Churlo, Chorrón, Richard, Román, María 7, Elenita, y por último Raulillo, al que le tengo que agradecer además una ayudita final con un "try-catch" que ya forma parte de esta tesis.

Y por último mi más sincero agradecimiento a mi familia: Mama, Papa, Juanjo, Carolina. Por vuestro apoyo y por ser como sois. Y cómo dejarme a las "nuevas incorporaciones", Zory y Alex, gracias a todos, nada de esto hubiera sido posible sin vosotros.

Alberto Pozo Esteban

Las Rozas de Madrid, Mayo del 2.016.

1

Introducción

DESDE hace décadas, el reconocimiento de actividades ha centrado una gran parte de los estudios científicos relativos a la inteligencia artificial, siendo las actividades humanas las que más atención han suscitado (aunque no son las únicas). Estos estudios han tratado de realizar una clasificación certera de las actividades de alto nivel que lleva a cabo un ser humano, (Weinland et al., 2011) y (Poppe, 2010), pudiendo partir de diferentes fuentes de información; como vídeo, GPS, RFID, cámaras TOC, etc. Sin embargo, la mayoría de los estudios centrados en el reconocimiento de actividades imponían la restricción de que hubiera un único elemento en la escena, o que al menos el reconocimiento fuera de la acción realizada por un único elemento. En los últimos años han surgido nuevos estudios que han tratado de eliminar dicha restricción, y se han centrado en tratar de realizar un reconocimiento de acciones más complejas, llevadas a cabo por un grupo de elementos que se coordinan persiguiendo un mismo fin. Este cambio de paradigma resulta natural si tenemos en cuenta que el ser humano es un animal social, y por lo tanto gran parte de las acciones que realiza son en realidad acciones de grupo. Existen multitud de ejemplos que constituyen dominios ideales para el reconocimiento de actividades de grupo; como la vigilancia, la defensa, los deportes de equipo, la etología, etc. Se trata por tanto de un campo relativamente novedoso, sobre el que faltan por asentar gran parte de los principios básicos tales como la representación más adecuada del problema, las características que deben utilizarse para la clasificación de las acciones, los métodos de aprendizaje, e incluso, los conjuntos de datos válidos para el aprendizaje y validación de los sistemas. La presente propuesta de tesis trata de afrontar algunas de estas cuestiones, proponiendo un novedoso modo de representación del problema, con una amplia cantidad de características propuestas, así como varios sistemas de selección de dichas características para una correcta y eficaz reducción de la dimensionalidad, y por último, un conjunto de datos que se adapta a las necesidades de

dicho campo de investigación.

1.1 Organización del documento

A lo largo del siguiente documento se hace un repaso del estado del arte actual en lo relativo al reconocimiento de actividades de grupo, así como se hace una propuesta novedosa tanto de un sistema de reconocimiento como de un conjunto de datos, con el fin de contribuir con ello a este campo de investigación, el documento se encuentra estructurado de la siguiente manera:

El capítulo 2 muestra de forma detallada el estado actual de los sistemas de reconocimiento de actividades, poniendo especial atención a los sistemas de reconocimiento de actividades de grupo. En primer lugar se muestra una clasificación de todos los tipos de sistemas, en función de diferentes criterios como el número de elementos, el nivel de fusión o el tipo de enfoque tomado, para más tarde detallar los trabajos más representativos de cada una de las áreas. Explicando en qué punto se encuentra actualmente el estado del arte, y dar una visión global de las diferentes alternativas existentes para la resolución del problema planteado.

El capítulo 3 recorre los diferentes conjuntos de datos disponibles para los sistemas de reconocimiento de actividades (poniendo especial atención a los sistemas de reconocimiento de actividades de grupo).

El capítulo 4 muestra las diferentes alternativas existentes en el campo de la selección de características, haciendo un resumen del estado del arte actual en dicho campo.

En el capítulo 5 se presenta una novedosa forma de modelar el problema general abordado en esta tesis, así como una implementación concreta para el dominio del 2 vs 2 en baloncesto. Aquí se detalla toda la estructura de las diferentes características utilizadas para modelar el problema así como todas las interrelaciones existentes.

El capítulo 6 muestra el sistema desarrollado a lo largo de esta tesis; empezando por la modelización inicial, describiendo después el tratamiento realizado con las características obtenidas y finalizando con el sistema de clasificación empleado. Por tanto, aquí se detalla una de las aportaciones principales de esta tesis, como un sistema de reconocimiento de actividades de grupo.

En el capítulo 7 se presenta un novedoso conjunto de datos desarrollado para el campo del reconocimiento de actividades de grupo. En él se describen todos los datos recopilados así como

la forma de interaccionar con ellos, tratando de ser una aportación valiosa en un campo de investigación donde no existen un gran número de conjuntos de datos adecuados y disponibles.

El capítulo 8 muestra los resultados obtenidos al aplicar el sistema descrito en el capítulo 6 sobre el conjunto de datos descrito en el capítulo 7. En el se detallan todos los resultados pormenorizados en función de las diferentes configuraciones iniciales del sistema, así como se describen los protocolos seguidos para la obtención de dichos resultados.

Por último, el capítulo 9 muestra toda una serie de conclusiones obtenidas a lo largo del desarrollo de esta tesis, así como un conjunto de ideas y propuestas para el desarrollo de trabajos futuros a partir de lo aquí presentado.

Además, en el apéndice A se han detallado los protocolos y métricas utilizados durante el transcurso de esta tesis para la obtención de resultados en los diferentes experimentos llevados a cabo.



Estado del Arte

EN este apartado se presentan los avances más significativos en el campo del reconocimiento de actividades en grupo. En primer lugar se realiza un repaso por las investigaciones que han obtenido un mayor éxito en los últimos años; posteriormente se hace especial hincapié en los conjuntos de datos existentes en la actualidad, ya que es un elemento crítico y fundamental en este tipo de investigaciones; por último se realiza un breve repaso a las técnicas de selección de atributos, por ser un punto importante en la investigación llevada a cabo.

2

Reconocimiento de actividades

El reconocimiento de actividades (fundamentalmente humanas, aunque también de otros tipos) ha sido uno de los campos más prolíficos en las últimas décadas. Esta actividad se encuentra relacionada con varias de las ramas de computación tales como el aprendizaje automático o la visión artificial. A fin de estructurar la gran cantidad de aportaciones existentes dentro de este campo de investigación, se muestra en la imagen 2.1 un resumen de los diferentes enfoques existentes en el amplio campo del reconocimiento de actividades, poniendo especial hincapié en el reconocimiento de actividades de grupo.

La primera aproximación lógica al reconocimiento de actividades se centra en reconocer comportamientos (más o menos complejos) de un determinado individuo. Existen en las dos últimas décadas gran cantidad de trabajos centrados en reconocer un conjunto concreto de actividades realizadas por un ser humano determinado, por ejemplo en (Cilla et al., 2011), se realiza el reconocimiento de la actividad llevada a cabo por un único individuo (en cada momento), a partir de los datos obtenidos por diferentes cámaras 2D (grabando desde diferentes puntos de vista), llegando a diferenciar entre diez comportamientos diferentes realizados por un total de nueve individuos, tal y como se presenta en el conjunto de datos Weizmann (wei, 2005). Existen otro muchos trabajos centrados en este mismo problema, tales como (Moeslund et al., 2006), (Aggarwal and Ryoo, 2011) o (Poppe, 2010). Sin embargo, es un hecho innegable que gran parte de los comportamientos complejos son ejecutados por conjuntos de individuos de forma coordinada. El ser humano, como animal social que es, supone un buen ejemplo de ello; como queda de manifiesto en muchas de las actividades humanas: deportes de equipo, maniobras militares, bailes, tráfico de vehículos, etc. Además, no solo el hombre realiza actividades de grupo de forma coordinada, estas acciones las podemos encontrar también en la naturaleza, la etología es la ciencia que estudia el comportamiento social de los animales.

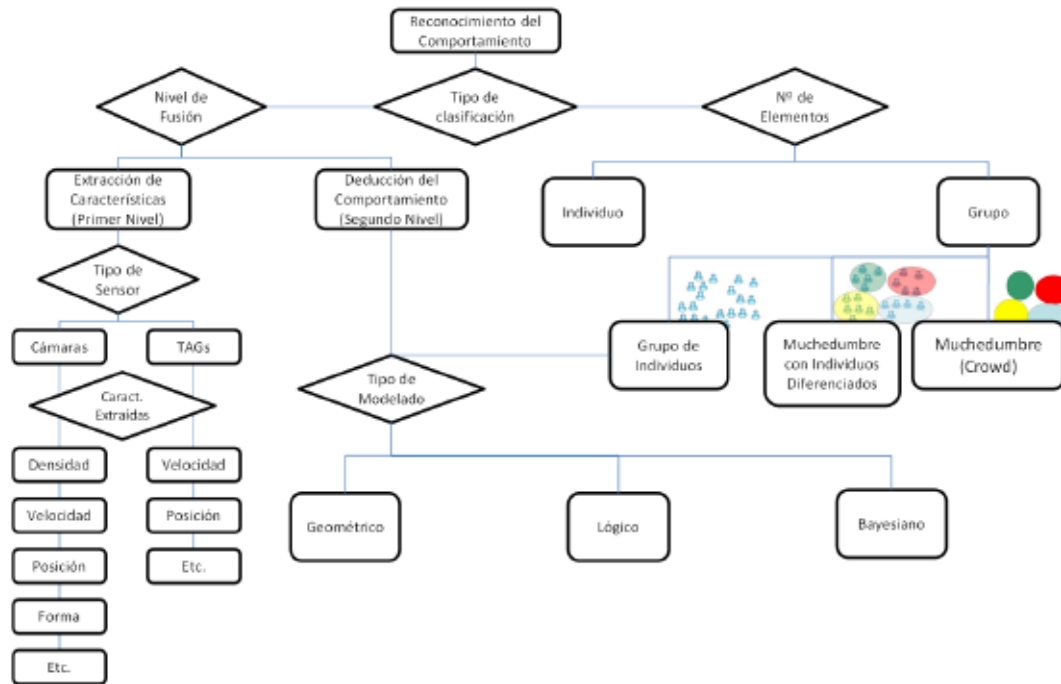


Figura 2.1: Clasificación del estado del arte en el reconocimiento de actividades

La mayoría de los trabajos sobre reconocimiento de actividades se encuentra dentro del campo de la visión artificial, donde se realiza un seguimiento ("tracking") del sujeto o sujetos a analizar. Al igual que en lo comentado anteriormente, en el seguimiento se comenzó por abordar el problema de un solo elemento, pero más tarde han ido apareciendo artículos en los que se hace el seguimiento de varios, artículos como (Yilmaz et al., 2006), (Li et al., 2013b) o (Yang et al., 2011) muestran un buen conjunto de ellos. Si bien en estos trabajos solo se realiza el seguimiento de los individuos, sin tratar de analizar comportamientos más complejos para obtener conocimientos sobre actividades de más alto nivel.

Sin embargo en los últimos años han empezado a aparecer propuestas interesantes relativas al reconocimiento de actividades en grupo. Uno de los primeros trabajos (Khan and Shah, 2005) en 2005, se basa en obtener las coordenadas u'_{fp} v'_{fp} de P puntos pertenecientes a cada uno de los miembros del grupo durante F fotogramas. Las coordenadas de dichos puntos son transformadas restandoles las coordenadas del punto medio de cada miembro del grupo. Así, las coordenadas u_{fp} v_{fp} equivalen a: $(u'_{fp} - ua_f, v'_{fp} - va_f)$ donde ua_f y va_f son la media de las coordenadas de todos los puntos. Con estos datos crea una matriz cuyo rango (en un caso perfecto sin nada de ruido) debe ser menor o igual a tres en caso de que los elementos del grupo mantengan una formación estructurada. Para evitar el problema del ruido se realiza una

factorización y minimización mediante un algoritmo iterativo.

En ese mismo año, en (Vaswani et al., 2005) se trata de reconocer no un comportamiento de grupo, sino una desviación del mismo. Sin embargo estos trabajos iniciales mantienen unas restricciones muy fuertes, siendo incapaces de realizar el reconocimiento de acciones más complejas en las que el comportamiento del grupo emana de los diferentes comportamientos individuales coordinados. Un ejemplo claro de estos comportamientos los podemos encontrar en los deportes de equipo. En ellos cada jugador lleva a cabo una acción individual diferente, y el comportamiento grupal coordinado emana del objetivo común que comparten todos los miembros del equipo.

Otro de los factores fundamentales que define la dificultad del problema en el reconocimiento de actividades en grupo radica en el número de elementos que lo componen. De esta manera, es lógico que muchas aproximaciones tratasen de abordar el problema empezando por grupos conformados por pocos individuos. Así, en (Zhou et al., 2008) mediante el seguimiento de dos elementos se clasifica el comportamiento de la pareja de entre cinco comportamientos posibles (perseguir, seguir, juntarse, independiente y caminar juntos).

En (Gong and Xiang, 2003) se trata de realizar el reconocimiento de comportamientos sencillos obtenidos del tracking de cuatro objetos diferentes en condiciones al aire libre, mediante la utilización de redes bayesianas. Para ello se realiza el seguimiento de cada objeto de la imagen y se le asigna a cada uno una "bounding box" (rectángulo mínimo de la imagen que contiene al objeto). Una vez obtenidas las diferentes "bounding box" se extrae el vector de características compuesto por: la coordenada x, la coordenada y, el ancho, el alto, el porcentaje de ocupación del objeto dentro de su "bounding box", la velocidad en el eje x y la velocidad en el eje y. Este vector de características es utilizado en el artículo como entrada para cuatro tipos diferentes de redes bayesianas dinámicas: "Multi-Observation Hidden Markov Model", "Parallel HMM", "Couple HMM" y "Dynamically Multi-Linked HMM". Tras comparar los resultados obtenidos por los diferentes sistemas se concluye que en caso de ejemplos de entrenamiento con cierto ruido, el DMLHMM obtiene los mejores resultados.

En general, los comportamientos de grupo pueden definirse como tal en la medida en la que todos los elementos del grupo comparten una meta común, y las acciones e interacciones que llevan a cabo se realizan con esa finalidad. Toda esta actividad tiene lugar a lo largo del tiempo, por lo que la componente temporal será un factor primordial en la mayoría de las propuestas. Por este motivo, una aproximación natural a este problema viene marcada por el

uso de redes bayesianas, como ocurre en el ya citado (Gong and Xiang, 2003) o (Zhang et al., 2006), (Hongeng and Nevatia, 2001), (Swears and Hoogs, 2009), (Hakeem and Shah, 2007), (Intille and Bobick, 2001), y (Liu and Chua, 2006). Nos encontraríamos por tanto en la hoja "bayesiano" del grafo representado en 2.1. En (Zhang et al., 2006) se presenta una arquitectura en dos niveles, en primer lugar realizan el reconocimiento de las actividades individuales, y en el segundo nivel categorizan las actividades de grupo. Las características utilizadas para el reconocimiento de actividades individuales son de tipo audio-visual (AV). Con ellas se realiza el reconocimiento de la actividad de cada uno de los individuos mediante un conjunto de HMMs (uno para cada una de las actividades individuales previamente escogidas). Una vez catalogada la acción que está llevando a cabo cada individuo, esta secuencia de acciones se pasa a la segunda capa (más las características audio-visuales en crudo no asociadas a ninguna actividad). En esta segunda capa se realiza el reconocimiento de la actividad de grupo, utilizando para ello un nuevo conjunto de HMMs. Tanto el conjunto de acciones individuales como el de acciones de grupo son finitos y definidos previamente. Por tanto solo se pueden identificar las acciones previamente definidas. Para el entrenamiento de los HMMs se utiliza un algoritmo de Expectation-Maximization (EM), y cada HMM se compone de una mezcla de gaussianas. Cuando aparece una nueva secuencia de entradas AV, se busca la secuencia de HMMs que encaja mejor (mediante un algoritmo de tipo Viterbi). Una vez obtenida la secuencia de HMMs se obtiene la secuencia de acciones individuales (ya que cada HMM corresponde con una acción individual) esta secuencia se pone como entrada a la segunda capa de HMM's junto con los datos AV no encajados en ninguna acción. Esta segunda capa es la encargada de decidir la acción de grupo llevada a cabo. El artículo presenta los resultados de reconocer entre nueve acciones de grupo diferentes entre las que se encuentran: Discusión, monólogo, Presentación (con proyector), Clase (con pizarra), etc.

Sin embargo no todos los trabajos se basan en redes bayesianas, si nos centramos en el nodo "lógico" del grafo 2.1 nos encontramos con (Perše et al., 2009) donde se hace un análisis de uno de los deportes de equipo más seguidos en todo el mundo, el baloncesto. En él, se realiza el estudio de los comportamientos en dos fases diferenciadas: la primera trata de segmentar el partido en secuencias de ataque, defensa y tiempo muerto, mediante el estudio de las velocidades de los jugadores. Para ello se compone el vector $x(t) = [x_t, y_t, Ax_t, Ay_t]$ que representa las posiciones del centroide de cada equipo en las coordenadas X e Y, y la velocidad de dicho centroide respectivamente. En la segunda fase, más compleja, se realiza el estudio detallado de las jugadas de ataque de uno de los equipos. Para ello se definen una serie de

comportamientos clave": bloqueo, movimiento y formación inicial. Se considera bloqueo a la acción por la cual un atacante sin balón y completamente quieto pone su cuerpo de manera que impida o dificulte el movimiento del defensor, es como una pared que hay que saltar si queremos seguir a nuestro atacante. Movimiento (tras dividir la cancha de baloncesto en zonas clave) se considera movimiento cada vez que un jugador abandona una de las zonas para situarse en una nueva, a fin de evitar considerar una cantidad abultada de movimientos en los casos en los que un jugador se mueva por la frontera entre dos zonas (más aún si consideramos el ruido de los datos de entrada) sólo se considerará que un jugador ha abandonado una zona para adentrarse en otra si lo hace por al menos medio metro de distancia. Por último, formación inicial se refiere a la zona en la que se encuentra cada uno de los jugadores del equipo al comienzo del ataque. Tras encontrar todos estos elementos clave, se realiza una construcción semántica con la que se construye un patrón. Si bien esta aproximación obtiene buenos resultados y conforma un avance valioso en el reconocimiento de actividades, puede resultar débil en el sentido de necesitar una gran cantidad de información proveniente del experto y resultar poco adaptable a diferentes dominios, debido a la gran dependencia existente entre los elementos clave y su dominio.

Otra aproximación no basada en redes bayesianas es (Morariu and Davis, 2011), donde se utilizan redes lógicas de Markov (Markov Logic Networks) con lógica de primer orden para analizar el comportamiento en el dominio del uno contra uno dentro del baloncesto. Realizando el seguimiento de la posición de las manos, los pies y el balón, se observan los eventos generados en base a esas trayectorias, utilizando esos eventos en la inferencia lógica posterior.

Por otro lado, dentro del tipo de modelado "geométrico"(2.1) autores como Chellappa y Ryoo disponen de varios artículos (Li et al., 2009), (Li and Chellappa, 2010), (Li et al., 2013a) en los que se trata de realizar el reconocimiento del comportamiento de un equipo de fútbol americano mediante el uso de tensores y matrices de interacción. En primer lugar, tratan de definir cada una de las interacciones ocurridas durante una secuencia mediante el tensor de cuatro dimensiones $Y(t_1, t_2, p_1, p_2)$, que indica la interacción del elemento p_1 en el instante t_1 con el elemento p_2 en el instante t_2 . Este tensor general, que en principio puede admitir información de muchos tipos, es construido mediante las trayectorias de cada uno de los jugadores en el campo; de esta manera, el tensor $Y(t_1, t_2, p_1, p_2)$ expresa la distancia entre el punto ocupado por el jugador p_1 en el instante t_1 y el jugador p_2 en el instante t_2 . Una vez construido, el tensor es reducido buscando un "kernel" tensor de rango 2, y dimensiones $P \times P$ que persigue tres objetivos concretos: independencia del punto de vista de la cámara,

máxima similitud entre elementos de la misma clase (al clasificar los posibles comportamientos) y máxima diferenciación entre elementos de clases distintas. Dicho kernel es empleado para reducir el tensor Y , obteniendo una representación matricial más compacta a la que han dado a llamar "discriminative temporal interaction matrix", utilizada para realizar la clasificación de los diferentes comportamientos de grupo. Estos trabajos han sido realizados utilizando el conjunto de datos "GaTech Football Plays Dataset", comentado más adelante. Basándose en este mismo conjunto de datos, en (Siddiquie et al., 2009) se realiza otra aproximación que arroja buenos resultados, basada en el uso de características espacio-temporales y de forma del grupo, mediante el aprendizaje basado en múltiples kernels (Gonen and Alpaydin, 2011).

Por último, en (Chen et al., 2015) tenemos una aproximación centrada en el deporte del baloncesto, en la que se trata de realizar un sistema muy poco supervisado de reconocimiento de la estrategia seguida por el equipo atacante. Para ello, parten de N vídeos compuestos cada uno por un ataque que sigue una de las estrategias del conjunto S . Obtienen de los vídeos las trayectorias de los jugadores (siempre de los atacantes), y comparan esas trayectorias midiendo las distancias que las separan (en grupos de dos) con la ayuda de un algoritmo de "Dynamic Time Warping". Una vez han calculado las distancias entre todas las trayectorias (corrigiendo las diferencias temporales con el DTW), miden las distancias que separan a cada vídeo, para lo cual emparejan las trayectorias de cada par de dos vídeos por similitud, y obtienen el sumatorio de dichas distancias. Tras obtener las distancias que separan a cada vídeo, realizan un cluster iterativo hasta obtener S clusters de vídeos, cada uno de los cuales representa alguna de las estrategias del conjunto inicial. Una vez que tienen esos clusters, utilizando de nuevo el DTW, realizan para cada cluster un nuevo proceso de clustering en el que tratan de clasificar todas las trayectorias en P clusters (siendo P el número de jugadores). De esta manera, tienen clasificadas para cada tipo de estrategia el conjunto de trayectorias llevadas a cabo por un jugador concreto. A partir de ahí obtienen la trayectoria "prototipo" de ese jugador para esa estrategia, mediante un nuevo ajuste DTW y haciendo regresión con un modelo Gaussiano. Finalmente seleccionan para cada cluster de vídeos el más característico (distancia mínima con la prototipo) y etiquetan la estrategia de ese vídeo mediante un experto. En este punto ya pueden clasificar cualquier nuevo vídeo simplemente calculando a que trayectorias prototipo se asemejan más las trayectorias del vídeo a clasificar.

Queda patente que nos encontramos ante un campo de la investigación en auge, con gran cantidad de dominios prácticos muy interesados en alentar y fomentar la investigación (como el deporte de élite o la seguridad) y con mucho camino por delante. Desde nuestro punto de vista

el deporte es el dominio perfecto para la investigación, tanto por contar con gran interés real como por su estructura reglada y controlada. Además, consideramos que los trabajos llevados a cabo hasta ahora, así como su análisis parecen indicar que un enfoque basado en la posición a lo largo del tiempo de los diferentes individuos del grupo puede ser muy adecuado para alcanzar la meta del reconocimiento de la actividad, poniendo especial énfasis en las características estudiadas.

3

Conjuntos de datos

Una parte fundamental en cualquier investigación consiste en disponer de un buen conjunto de datos con el que hacer los experimentos. En la mayoría de ocasiones ya existen diferentes conjuntos de datos preparados para ser utilizados en las diferentes ramas de la ciencia, sin embargo en otras ocasiones el investigador debe comenzar por crear su propio conjunto de datos y trabajar con él. Tal y como se comentó en el apartado anterior el reconocimiento de actividades individuales ha ocupado gran parte de la investigación en las últimas décadas, es por tanto lógico que existan gran variedad de conjuntos de datos que pueden ser utilizados con este propósito. Como ejemplos más característicos podemos encontrar KTH, Weizmann e Ixmas.

- KTH: Se compone de una secuencia de vídeo con seis tipos de acciones individuales diferentes (andar, trotar, correr, boxear, saludar y aplaudir). Cada acción es repetida varias veces por distintos individuos en cuatro escenarios diferentes. Puede encontrarse aquí: (kth, 2004)
- Weizmann: Contiene diez tipos de acciones diferentes (correr, andar, esquivar, saltar repetidamente, gatear, saltar hacia delante con los pies juntos, saltar en vertical con los pies juntos, saludar con dos manos, saludar con una mano e inclinarse). Cada acción es repetida una sola vez por cada uno de los nueve actores, dando un total de nueve ejemplos de cada acción. Puede encontrarse en: (wei, 2005)
- Ixmas: Incluye trece tipos de acciones (mirar el reloj, cruzar los brazos, mover la cabeza, sentarse, levantarse, dar una vuelta, andar, saludar, dar un puñetazo, dar una patada, señalar, coger un objeto y lanzar un objeto). Cada acción es repetida tres veces por cada

uno de los 11 actores, y se dispone vídeo desde cinco puntos de vista diferentes. Puede encontrarse en: (ixm, 2006)

Cada uno de estos conjuntos de datos ha sido utilizado por múltiples trabajos centrados en el reconocimiento de actividades humanas individuales. Existen además muchos otros, sin embargo, al dar el salto al reconocimiento de actividades en grupos, el conjunto de datos disponible se reduce considerablemente. En (Zhou et al., 2008), los autores crean su propio conjunto de datos para poder probar y validar su sistema. Este conjunto de datos contiene escenas con dos participantes, que realizan una de las siguientes cinco acciones: perseguirse (uno persigue y el otro trata de huir), seguirse (uno se mueve libremente, y el otro le sigue), movimiento independiente (ambos se mueven libremente con independencia de lo que haga el otro), reunión (ambos confluyen en un punto acordado) y caminar juntos. El conjunto de datos se compone de 170 instancias repartidas entre estas cinco acciones. Una de las limitaciones más evidentes de este conjunto de datos consiste en que todas las acciones son realizadas únicamente por dos individuos, lo que puede resultar escaso para comprobar algunos sistemas. Otro de los conjuntos de datos más utilizados en el reconocimiento de actividades en grupo (Li et al., 2009), (Li et al., 2013a), (Li and Chellappa, 2010) o (Siddiquie et al., 2009) es el GaTech Football Dataset. Este conjunto de datos se compone de 155 vídeos de fútbol americano (NCAA). Cada video corresponde con una jugada de fútbol americano, y se dispone también de las anotaciones relativas al tipo de jugada ofensiva que se lleva a cabo, así como de las posiciones de los jugadores en cada momento. Estas características le convierten en uno de los conjuntos de datos más valioso dentro del campo del reconocimiento de actividades de grupo. Por último, otro de los conjuntos de datos más apropiado puede encontrarse en (Perš, 2005), y se llama CVBASE'06. Este conjunto de datos creado en la universidad de Ljubljana, por los autores Janez Perš, Marta Bon y Goran Vuckovic está compuesto por tres grupos diferentes de vídeos correspondientes a tres deportes diferentes: baloncesto, balonmano y squash. Además de un conjunto de funciones Matlab que permiten el acceso a la información. A continuación se realiza un análisis de los datos proporcionados para el deporte: Balonmano, ya que es del que se dispone de la información necesaria para ser utilizado en el análisis del comportamiento en grupos. Para el balonmano se dispone de tres vídeos correspondientes a un mismo partido de diez minutos de duración. Cada uno de los vídeos corresponde con una cámara diferente, dos de ellas situadas de manera fija en posición cenital (A y B), y una tercera situada en una banda enfocando siempre al centro de la jugada. Cada uno de los tres vídeos se encuentra sincronizado con los demás fotograma a fotograma. Además de los vídeos suministrados se

dispone de seis archivos de texto plano (.txt) en los que se incluye información sobre la posición y la acción desarrollada por cada jugador, así como de la acción conjunta desarrollada por el equipo. A continuación se presenta de manera detallada la información contenida en cada uno de los archivos de texto:

- **PlayerList.txt:** Nombre y número de cada uno de los jugadores. El nombre representa el color de la camiseta que se observa en el vídeo.
- **Positions.txt:** Posición de cada uno de los jugadores en cada uno de los fotogramas del vídeo. Cada jugador se representa por su nombre, y la posición se muestra en dos dimensiones representada en tres ejes de coordenadas diferentes: El de la pista, y el de cada una de las dos cámaras cenitales. Además se indica la información sobre qué cámara se ha realizado el posicionamiento.
- **TeamActivityDictionary.txt:** Proporciona la traducción entre las anotaciones hechas por el experto (un entrenador de balonmano) y su significado para cada una de las posibles acciones llevadas a cabo por un equipo. Además se asocia un número a cada una de las posibles acciones, dicho número será utilizado por diferentes funciones de Matlab.
- **PlayerActivityDictionary:** Proporciona la traducción entre las anotaciones hechas por el experto (un entrenador de balonmano) y su significado para cada una de las posibles acciones individuales de un jugador. Además se asocia un número a cada una de las posibles acciones, dicho número será utilizado por diferentes funciones de Matlab.
- **PlayerActivity.txt:** Cada una de las acciones llevadas a cabo por cada jugador. Se supone que cada actividad es producida en un instante determinado con duración nula. La información contenida incluye el instante en que ha sido observada la actividad, qué actividad ha sido observada, y qué jugador la ha llevado a cabo.
- **TeamActivity.txt:** Cada una de las acciones llevadas a cabo por el equipo. Cada acción tiene un instante de inicio y un instante final (por lo que las acciones, al contrario que las individuales, tienen una duración). Es importante destacar que cada acción empieza en el instante en que terminó la acción anterior, por lo que el equipo se encuentra en todo momento llevando a cabo alguna de estas acciones. Además de la información detallada anteriormente, se dispone de un conjunto de funciones Matlab preparadas para proporcionar el acceso a los datos de manera rápida y cómoda.

Nos encontramos por tanto ante una de las debilidades de este campo de investigación, la escasez de conjuntos de datos válidos para el desarrollo y la evaluación de las diferentes técnicas llevadas a cabo. Dado que uno de los puntos fundamentales en la investigación consiste en la comparación de las técnicas desarrolladas con las anteriormente existentes en el estado del arte, es importante que los experimentos llevados a cabo se realicen con conjuntos de datos de acceso público, con los que cualquiera pueda probar sus técnicas y realizar así la comparación. Sin embargo, es habitual dentro del reconocimiento de actividades en grupo encontrarnos con trabajos de gran calidad, que sin embargo realizan sus experimentos con conjuntos de datos no públicos, por lo que la presencia de más conjuntos de datos es vital, y aun más importante es que estos tengan un acceso libre para que cualquier investigación pueda acceder a ellos. Estos conjuntos de datos deben aportar dos tipos de información básica para ser adecuados para esta rama de la investigación. Por un lado, información del tipo de actividad de grupo llevada a cabo, y por otro, información relativa a los individuos que la llevan a cabo, a partir de la cual se puedan desarrollar los sistemas pertinentes.

4

Selección de características

Otro de los apartados importantes en el campo del reconocimiento de actividades en grupo es la selección de características. Esto se debe a que al aumentar el número de elementos en la escena (respecto al reconocimiento de actividades individuales) el número de características que se pueden extraer también aumenta. Por tanto, en muchas ocasiones puede ser aconsejable realizar una selección previa de las características más representativas de la acción. Esta selección puede ser mediante la utilización de un experto, o mediante mecanismos automáticos ideados con tal fin. A continuación se presenta un breve resumen de las técnicas automáticas existentes para la selección de características.

De forma general podemos definir los algoritmos de selección de características como aquellos que de entre el total de las características seleccionan el subconjunto que optimiza la evaluación, siendo ese subconjunto:

- De un número determinado de elementos.
- Del menor número de elementos posible.
- Del tamaño más óptimo en función de la medida de evaluación obtenida.

Con el objetivo final de optimizar el aprendizaje, tanto en la velocidad de los algoritmos como en su capacidad para generalizar o en la simplicidad de su representación.

El objetivo de cualquier algoritmo de selección puede resumirse en una búsqueda de la relevancia de cada característica, si bien no existe una definición fija de relevancia (Bell and Wang, 2000), en la literatura actual existen gran cantidad de ellas, partamos de la idea de que sea E_i (con $1 \leq i \leq n$) el conjunto de los dominios de las características $X = x_1, \dots, x_n$.

Consideremos p una función de probabilidad de E sobre T , siendo el conjunto de las posibles clases (de un problema de clasificación). Entonces el objetivo consiste en encontrar un modelo c tal que $c: E \rightarrow T$ si utilizamos el conjunto adecuado de características. Si disponemos de un conjunto S de muestras de la función p . Entonces podemos tener:

- Relevancia respecto al objetivo: La característica $x_i \in X$ es relevante respecto al objetivo de clasificación c si existen dos instancias A, B en el espacio E tal que A y B sólo se diferencian en el valor de x_i y además tienen clases diferentes (Blum and Langley, 1997).
- Relevancia respecto al conjunto de muestras: La definición es similar para x_i salvo que en este caso se hace respecto al conjunto de muestras S , es decir, dada la característica $x_i \in X$ será relevante si existen dos instancias $A, B \in S$ tal que sólo se diferencian en la característica x_i y además tienen diferente clase.
- Relevancia respecto a la función p : La característica $x_i \in X$ es relevante respecto a la función p si existen dos instancias A, B en el espacio E tal que solo se diferencian en la característica x_i y $p(A) \neq 0$ y $p(B) \neq 0$ y tienen clases diferentes.
- Relevancia en función de mínima redundancia y máxima relevancia: Tal y como se muestra en (Peng et al., 2005) este algoritmo evalúa la información aportada por cada característica.
- Utilidad incremental (Gärdenfors, 1978): Dado un conjunto de características X' , la relevancia de x_i se mide como la diferencia entre la precisión del resultado obtenido con el algoritmo L usando el conjunto unión entre X' y x_i , y la precisión obtenida sólo con el conjunto X' .

Además de las diferentes posibles definiciones de relevancia, los diferentes algoritmos de selección de características se diferencian en la manera en la que van guiando la búsqueda entre los posibles subconjuntos de características. Por un lado deben centrarse en la manera en la que van modificando cada subconjunto y por otro en la manera en la que recorren las diferentes posibilidades.

Si nos centramos en la manera en la que van generando los diferentes subconjuntos de características debemos diferenciar entre dos grupos bien diferenciados:

- Crecientes: En los sistemas crecientes como (Maldonado and Weber, 2009) o (Patricio and Maravall, 2000), se parte de un conjunto vacío de características, y se van añadiendo

aquellas que mayor relevancia aportan, hasta llegar a un número previamente establecido (en caso de que exista) o hasta que no se consiga aportar más relevancia de la ya existente con el conjunto anterior.

- **Decrecientes:** En estos sistemas, por el contrario, partimos del conjunto total de características y vamos eliminando aquellas que menos relevancia aporten (Koller and Sahami, 1995).

Por otro lado, centrándonos en el tipo de búsqueda realizado, podemos diferenciar entre:

- **Exponencial:** Corresponden a algoritmos con un coste computacional de $O(2^n)$. De ello resulta una búsqueda óptima, que garantiza encontrar la mejor opción, teniendo que llegar hasta el final sólo en los casos en los que la función de evaluación no sea monótona (Narendra and Fukunaga, 1977). Otro ejemplo es el algoritmo A* con una heurística válida (Pearl, 1984).
- **Secuencial:** En este tipo de búsqueda se selecciona una de los posibles sucesores en cada iteración, y una vez que se fija una de las opciones no es posible volver atrás. Obteniendo así una complejidad computacional polinomial de $O(n^{k+1})$ siendo k el número de subconjuntos evaluado (iteraciones) en cada cambio.
- **Aleatoria:** Con este tipo de búsqueda se persigue que la aleatoriedad evite que el algoritmo caiga en un mínimo local pero esto también provoca que puedan ocurrir pasos en los que se empeora la relevancia respecto al paso anterior. Estos sistemas pueden dar varios subconjuntos finales de soluciones (Liu and Motoda, 2012).

En función de la relación que exista entre el algoritmo de sección de características y el algoritmo de clasificación utilizado podemos establecer una categoría de tres niveles entre los que se distinguen:

- **Embebidos:** Aquellos en los que el propio algoritmo de clasificación ya cuenta con algún mecanismo de selección de características (o ponderamiento) interno. Dentro de esta categoría se encuentran los métodos de inducción de conjunciones lógicas (Vere, 1975), (Winston, 1970) o las redes de neuronas (Mitchell, 1982).

- **Filtros:** En estos casos la selección de características se realiza antes (y con independencia) del algoritmo de clasificación usado posteriormente. Existen multitud de ejemplos como (Fish et al., 2012), (Peng et al., 2005) o (Liu et al., 2009).
- **Wrappers:** En este caso son los algoritmos de filtrado de características los que utilizan al algoritmo de clasificación para evaluar la relevancia de un determinado subconjunto de características (John et al., 1994). Es obvio destacar la idoneidad de este esquema, si bien cuenta con la desventaja de suponer en algunos casos un coste computacional muy alto al tener que pasar el algoritmo de clasificación para cada uno de los subconjuntos de características generados.

A lo largo de la propuesta realizada se estudiarán diferentes algoritmos de selección de características evaluando posibles enfoques para la relevancia así como diferentes alternativas para la relación entre el algoritmos de selección y el de clasificación.



Propuesta

ESTA tesis trata de realizar tres aportaciones principales al campo del reconocimiento de actividades en grupo. En primer lugar una novedosa modelización del problema de las actividades de grupo, basada principalmente (aunque no en exclusiva) en la localización de sus integrantes. En segundo lugar un nuevo sistema de reconocimiento de actividades en grupo, basado en la representación anterior, y con una fuerte influencia de la selección de las características más representativas. En último lugar, se ha desarrollado un nuevo conjunto de datos adecuado para el reconocimiento de actividades en grupo, por tratarse de un campo de investigación con un escaso número de conjuntos de datos. Los próximos capítulos describen con mayor detalle cada una de estas contribuciones. En primer lugar, en el apartado "Modelado y Diseño del Reconocimiento de Actividades de Grupo" se explican los detalles del modelado del problema, así como su contexto y características; más adelante, en el apartado "Sistemas de Reconocimiento De Actividades de Grupo" se describe el sistema desarrollado, especificando el funcionamiento tanto de la obtención como la selección de características, así como los métodos de reconocimiento implementados; por último, en el siguiente apartado "Inef12Basketball Dataset" se detalla el conjunto de datos realizado.

5

Modelado y Diseño del Reconocimiento de Actividades

El reconocimiento de actividades en grupo es un novedoso campo de investigación que resulta de eliminar la restricción de un solo individuo en el reconocimiento de actividades. La necesidad de eliminar esta restricción surge de forma natural cuando observamos que gran parte de las acciones (tanto humanas, como de otra índole) tienen una componente social y por tanto debe ser evaluado el grupo completo, y no cada uno de sus componentes individualmente. Este campo de investigación tiene gran cantidad de potenciales dominios de acción, tales como los deportes de equipo, la defensa, la etología, la vigilancia automática, etc.

Una de las características fundamentales en el reconocimiento de actividades es que estas se producen a lo largo de un determinado tiempo. Por lo tanto es una secuencia de observaciones (en lugar de una aislada) lo que debe ser clasificado, teniendo en cuenta que cada observación tiene una alta dependencia de la anterior. De esta característica fundamental en el reconocimiento de actividades surge la idea de dividir el problema en dos tipos: clasificación y segmentación. En el primer caso disponemos de un conjunto de secuencias segmentadas, donde en cada una de las secuencias se realiza una única y completa acción, que debe ser clasificada. En el segundo caso disponemos de una secuencia continua con varias acciones dentro, esta secuencia debe ser segmentada para poder llegar a la situación anterior. En esta propuesta vamos a centrarnos en el primer problema, siendo el segundo problema uno de los posibles trabajos futuros a llevar a cabo.

5.1 Descripción General del Problema

Tal y como se ha comentado, el reconocimiento de actividades de grupo puede emplearse en gran número de dominios diferentes. Sin embargo, se pueden extrapolar una serie de características comunes a todos los dominios. De forma general, podemos describir cualquier dominio de la siguiente manera: Existe una escena S , compuesta por un número M de grupos y un conjunto F de características relativas a la escena (eq:5.1). Cada grupo, puede desarrollar una actividad de un conjunto B de actividades (eq:5.2). Cada característica y cada comportamiento se produce a lo largo del tiempo.

$$S_t = \{G_t^1, G_t^2, \dots, G_t^M, F_t^1, F_t^2, \dots, F_t^{O_S}, B\} \quad (5.1)$$

$$|1 \leq t \leq T$$

$$B = \{b_1, b_2, \dots, B_p\} \quad (5.2)$$

Cada grupo está compuesto por un número N_m de elementos así como por una cantidad O_G de características propias del grupo y dependientes del dominio del problema (eq:5.3). Además, cada grupo está llevando a cabo un comportamiento del conjunto B (eq: 5.2) en cada instante de tiempo.

$$G_t = \{I_t^1, I_t^2, \dots, I_t^{N_m}, H_t^1, H_t^2, \dots, H_t^{O_G}\} | 1 \leq t \leq T$$

Finalmente, cada elemento se compone de un número O_I de características propias (también dependientes del dominio del problema). Estas características van cambiando su valor a lo largo del tiempo, por lo que en el instante t , un elemento n está definido por la ecuación 5.3.

$$I_t^n = \{J_t^1, J_t^2, \dots, J_t^{O_I}\} | 1 \leq t \leq T \quad (5.3)$$

Por lo tanto, una escena se compone de M grupos y O_S características, cada grupo lo forman N_M elementos y O_G características, y por último, cada elemento tiene O_I características. Quedando así una jerarquía de características relativas a la escena, los grupos y los individuos. Esta jerarquía resulta muy útil a la hora de clasificar los diferentes tipos de características, tanto para el sistema, como para la coordinación con el experto del dominio consultado. Finalmente, aunque todas las características son dependientes del dominio, pueden clasificarse en cuatro grupos principales:

- Real: Representa una característica que puede tener cualquier valor real, como por ejemplo la distancia a un punto. (eq: 5.4).

$$p_{nf}|f = 1 \dots F, n = 1 \dots N, p \in (-\infty, \infty) \quad (5.4)$$

- Boleana: Representa cualquier característica binaria cuyo valor puede ser 1 ó 0. Un ejemplo sería si un jugador posee el balón (en algún deporte de equipo) o no en un instante dado. (eq: 5.5).

$$b_{nf}|f = 1 \dots F, n = 1 \dots N, b = \{0, 1\} \quad (5.5)$$

- Relativa: Característica que indica el porcentaje de cumplimiento de una determinada situación. Puede tomar cualquier valor real entre 0 y 1. Por ejemplo, puede ser la intensidad lumínica percibida por un sensor. (eq: 5.6).

$$r_{nf}|f = 1 \dots F, n = 1 \dots N, r = [0, 1] \quad (5.6)$$

- Enumerador: Cualquier característica que deba representar un número finito de posibles valores. Por ejemplo es útil para representar los diferentes roles dentro de un grupo. (eq: 5.7).

$$e_{nf}|f = 1 \dots F, n = 1 \dots N, e = \{1, 2, 3, \dots, K\} \\ |K = \text{"possible values of feature"} \quad (5.7)$$

Qué conjunto de características se obtienen de una escena es una decisión muy importante que puede en la mayoría de ocasiones ser dependiente del dominio. Sin embargo, la gran mayoría de dominios comparten las características relativas a la posición de los elementos que componen los grupos, y por tanto estas características son tomadas como base para el modelado aquí mostrado.

Esta representación puede expresar cualquier escena en el que se lleve a cabo un comportamiento de grupo, tanto para el problema de la clasificación como para el de la segmentación. No obstante, en esta aproximación el esfuerzo va a centrarse en los problemas de clasificación, en los que se dispone de una escena segmentada, en la que se lleva a cabo un comportamiento grupal en cada segmento de la escena. Por lo que en la fase de entrenamiento el sistema recibirá

un conjunto de secuencias S_1, S_2 , etc. Etiquetadas cada una de ellas con un comportamiento concreto b_p del conjunto de posibles comportamientos B (eq:5.2). Una vez realizado el entrenamiento, una nueva secuencia sin etiquetar S_j será pasada al sistema, y este deberá realizar la clasificación indicando qué comportamiento grupal se está llevando a cabo en dicha secuencia.

Dado que se trata de una aproximación general, esta representación es válida para la inmensa mayoría de los problemas propuestos en el apartado 2, tanto los relacionados con el baloncesto ((Perše et al., 2009), (Morariu and Davis, 2011)) o con el fútbol americano ((Li et al., 2009), (Li and Chellappa, 2010)), como los que se ocupan de otros dominios diferentes ((Khan and Shah, 2005), (Vaswani et al., 2005), (Zhou et al., 2008)). Ya que en todos ellos es posible definir la estructura de características pertenecientes a los individuos, los grupos y la escena, si bien en algún caso alguno de los niveles puede quedar desierto.

5.2 Modelado del Dominio 2 vs 2 en Baloncesto

Uno de los dominios en los que más se está trabajando dentro del campo de reconocimiento de actividades de grupo es el de los deportes de equipo. Esto es debido fundamentalmente a dos motivos: en primer lugar por tratarse de un dominio que encaja a la perfección dentro del estudio de las actividades de grupo, por disponer de gran cantidad de acciones coordinadas, dentro de un marco de reglas bien conocido; en segundo lugar, por ser un campo con gran interés socio-económico, y disponer de la capacidad económica necesaria para realizar grandes inversiones en infraestructura.

Sin embargo, en muchas ocasiones el número de participantes en un deporte de equipo es muy elevado, y puede resultar un reto demasiado ambicioso en los primeros pasos de un sistema novedoso. Por todo ello, en esta propuesta de tesis hemos tratado de centrar los esfuerzos en el dominio concreto de dos contra dos dentro del deporte del baloncesto. Consideramos que es un dominio idóneo para crear y probar nuestro sistema, por tratarse de acciones claramente coordinadas, con un número reducido de elementos (cuatro, en total), en un ámbito de reglas bien definidas. También es importante destacar que si bien este dominio es una simplificación del dominio original del baloncesto, gran parte de lo aprendido en el puede ser extrapolable ya que en muchas ocasiones las acciones de grupo dentro de un equipo no son llevadas a cabo por todos los integrantes, sino que es realizada por un subconjunto menor de miembros del equipo.

Por todo ello se ha contado con el conjunto de datos (Pozo et al., 2012a) (descrito

en la siguiente sección) en el dominio del dos contra dos de baloncesto, que proporciona información detallada de 23 secuencias, cada una de las cuales representa una repetición de entre cinco acciones de grupo diferentes. En todas las secuencias la pareja atacante realiza la misma acción (movimiento "pick and roll") mientras que la pareja que defiende lleva a cabo diferentes comportamientos (defensas ante el "pick and roll"). Es importante destacar, que si bien el conjunto de datos dispone de ocho tipos de defensas diferentes, se ha seleccionado un subconjunto de estas por ser las que más repeticiones contenían. Por tanto, se dispone de 23 repeticiones de entre cinco posibles acciones de grupo (la defensa llevaba a cabo ante el ataque "pick and roll", es decir: "pasar por delante", "pasar de tercero", "push", cambio 2 "dos contra uno"). El conjunto de datos proporciona información sobre las posiciones en X e Y de cada uno de los jugadores (relativas a la esquina inferior izquierda de la cancha) con características de tipo real (ver: 5.4) y la información de quién lleva el balón en cada momento con una característica de tipo enumerador (ver: 5.7). Además de la información de la acción de grupo llevada a cabo en cada secuencia.

Partiendo de esta información básica, se ha procedido a modelar el problema obteniendo un conjunto total de cuarenta características, 39 reales y una enumerada. Estas características describen la posición (2D) de cada uno de los jugadores, la posición del centro de masas de cada uno de los grupos (cada una de las parejas, así como el formado por los cuatro jugadores), la posición de cada uno de los jugadores relativa al jugador que lleva el balón, la velocidad de cada uno de los jugadores, la velocidad del centro de masas de cada uno de los grupos, y por último, tres componentes del invariante laplaciano del grafo formado por los cuatro jugadores. Además de la característica enumerada que indica el jugador que lleva el balón en cada momento. Todas estas características pueden ser observadas en la tabla 5.1, además, las siguientes imágenes 5.1, 5.2, 5.3, 5.4 proporcionan una explicación gráfica de algunas de ellas.

Se dispone por tanto de una escena con cuatro individuos ($N = 4$), distribuidos en dos grupos ($M = 2$) de manera homogénea ($N_1 = 2$ y $N_2 = 2$). Dicha escena dispone de ocho características ($O_S = 8$) relativas a la posición y velocidad del centro de masas del conjunto de los jugadores, el dueño del balón y las tres componentes del invariante laplaciano del grafo formado por los cuatro jugadores. Cada elemento tiene seis características ($O_I = 6$) relativas a su posición, su velocidad y su posición relativa al dueño del balón. Por último, cada grupo dispone de cuatro características ($O_g = 4$) que indican su posición y su velocidad. Una vez obtenidas todas las características partiendo de los datos proporcionados por (Pozo et al.,

#	Característica
1	Coordenada X del jugador 1
2	Coordenada Y del jugador 1
3	Coordenada X del jugador 2
4	Coordenada Y del jugador 2
5	Coordenada X del jugador 3
6	Coordenada Y del jugador 3
7	Coordenada X del jugador 4
8	Coordenada Y del jugador 4
9	Coordenada X del centro de masas del equipo atacante
10	Coordenada Y del centro de masas del equipo atacante
11	Coordenada X del centro de masas del equipo defensor
12	Coordenada Y del centro de masas del equipo defensor
13	Coordenada X del centro de masas de todos los jugadores
14	Coordenada Y del centro de masas de todos los jugadores
15	Dueño actual del balón
16	Coordenada X del jugador 1 respecto al dueño del balón
17	Coordenada Y del jugador 1 respecto al dueño del balón
18	Coordenada X del jugador 2 respecto al dueño del balón
19	Coordenada Y del jugador 2 respecto al dueño del balón
20	Coordenada X del jugador 3 respecto al dueño del balón
21	Coordenada Y del jugador 3 respecto al dueño del balón
22	Coordenada X del jugador 4 respecto al dueño del balón
23	Coordenada Y del jugador 4 respecto al dueño del balón
24	Velocidad en el eje X del jugador 1
25	Velocidad en el eje Y del jugador 1
26	Velocidad en el eje X del jugador 2
27	Velocidad en el eje Y del jugador 2
28	Velocidad en el eje X del jugador 3
29	Velocidad en el eje Y del jugador 3
30	Velocidad en el eje X del jugador 4
31	Velocidad en el eje Y del jugador 4
32	Velocidad en el eje X del equipo atacante
33	Velocidad en el eje Y del equipo atacante
34	Velocidad en el eje X del equipo defensor
35	Velocidad en el eje Y del equipo defensor
36	Velocidad en el eje X del conjunto de todos los jugadores
37	Velocidad en el eje Y del conjunto de todos los jugadores
38	Primera componente laplaciana del grafo de los jugadores
39	Segunda componente laplaciana del grafo de los jugadores
40	Tercera componente laplaciana del grafo de los jugadores

Cuadro 5.1: Lista de Características

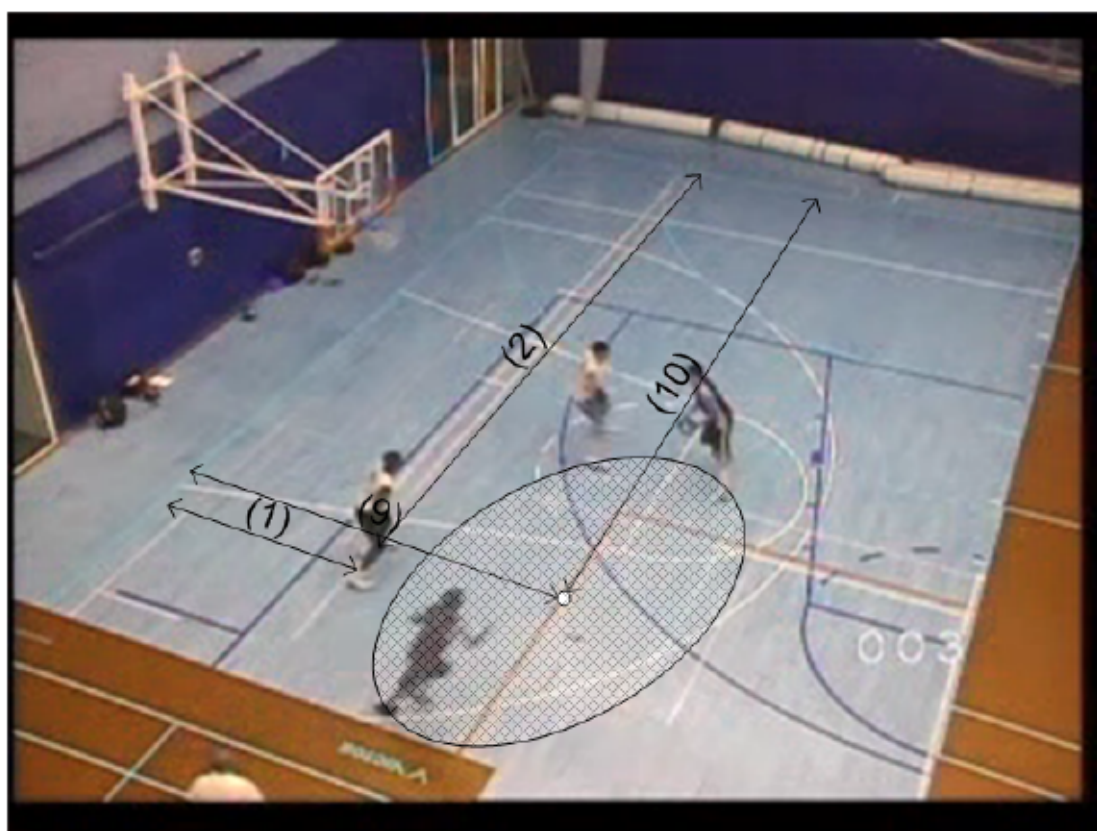


Figura 5.1: Coordenadas X e Y del equipo atacante

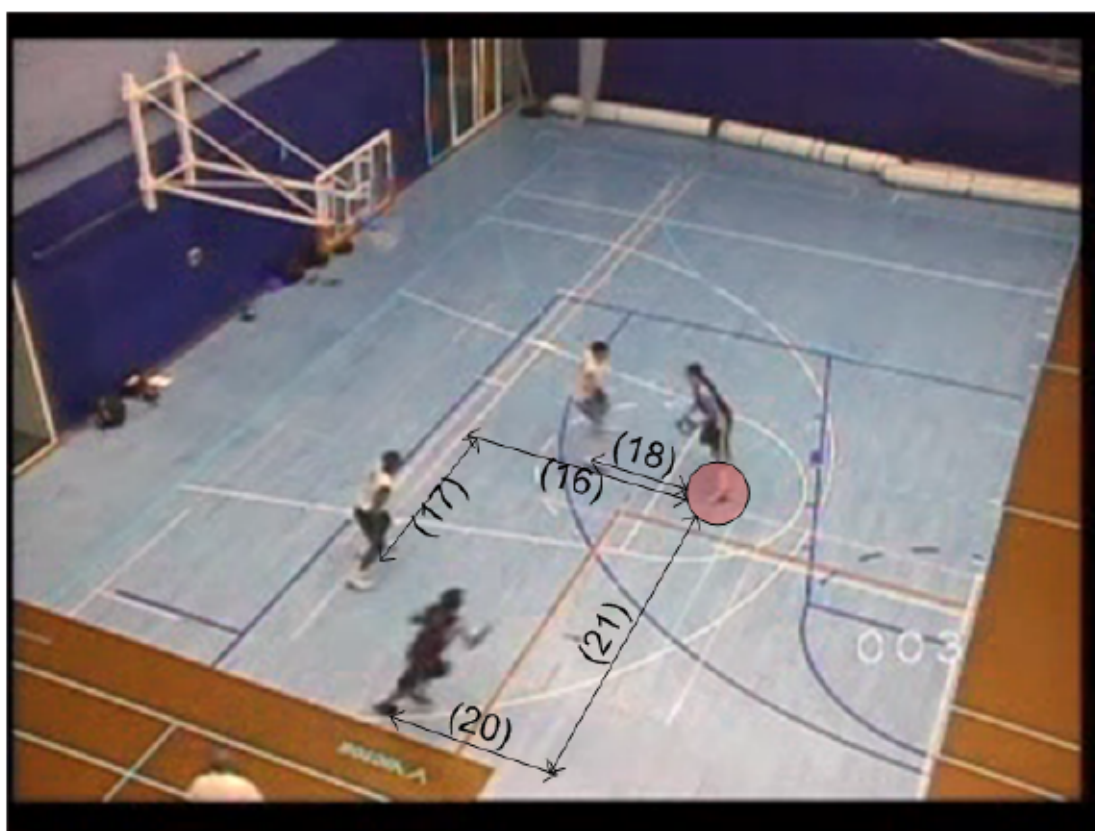


Figura 5.2: Coordenadas X e Y del resto de jugadores respecto del que lleva el balón

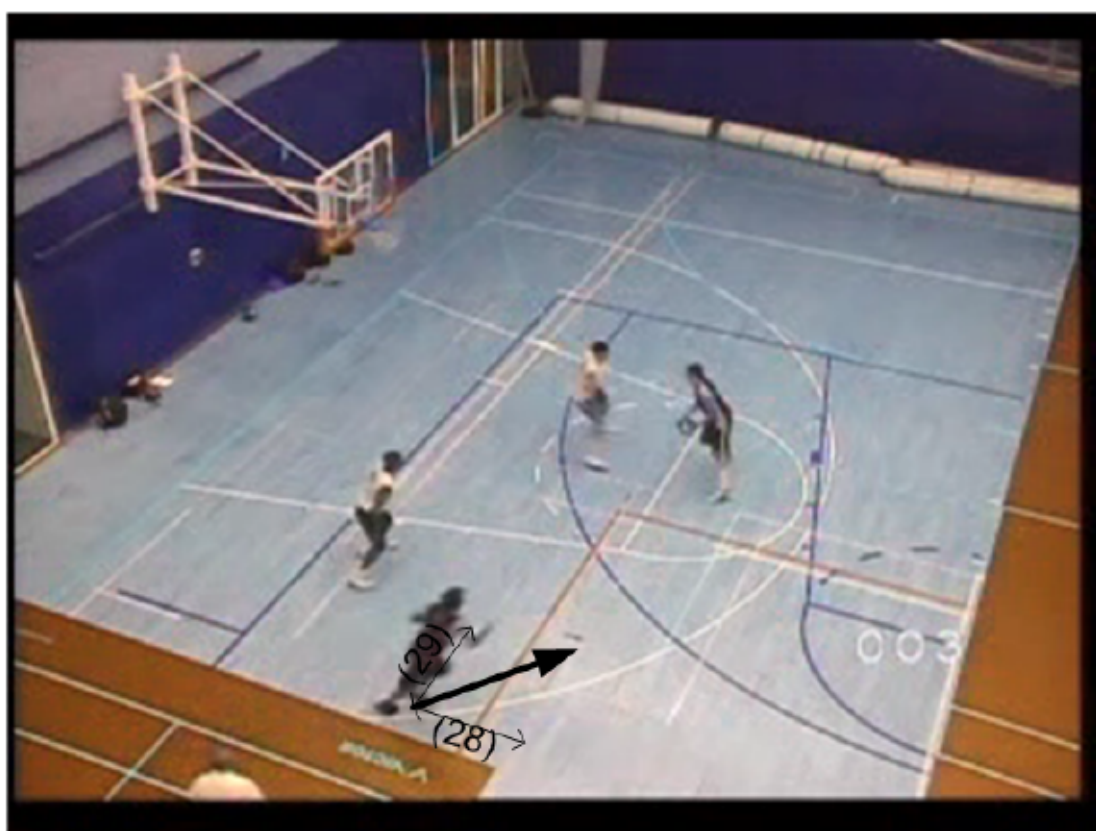


Figura 5.3: Coordenadas X e Y de la velocidad de un jugador

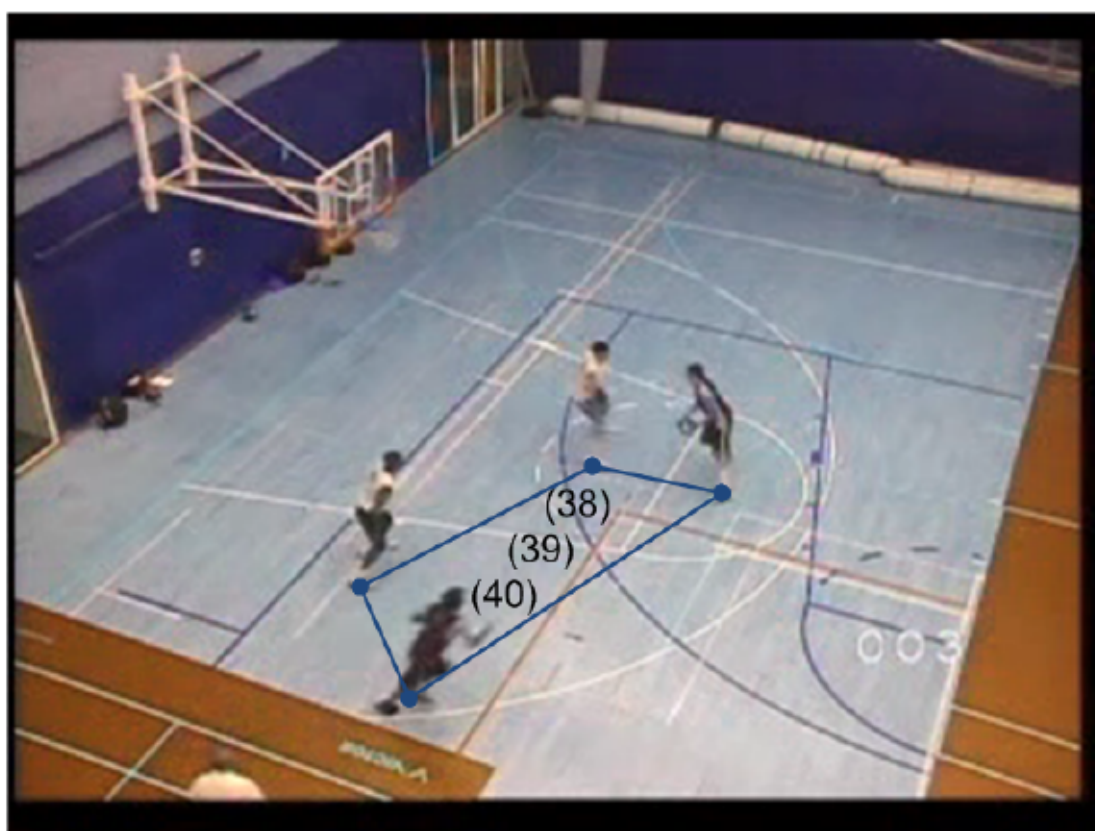


Figura 5.4: Componentes laplacianos del grafo formado por los cuatro jugadores

2012a) se dispone de la representación mostrada en 5.8.

$$\begin{aligned}
 S_t &= \{I_t^1, I_t^2, I_t^3, I_t^4, G_t^1, G_t^2, F_t\} \text{ con} \\
 I_t^n &= \{X_t^n, Y_t^n, XX_t^n, YY_t^n, V_{X_t^n}, V_{Y_t^n}\}, \\
 G_t^m &= \{X_t^m, Y_t^m, V_{X_t^m}, V_{Y_t^m}\}, \\
 F_t &= \{X_t, Y_t, V_{X_t}, V_{Y_t}, H_t, L_t^1, L_t^2, L_t^3\}
 \end{aligned} \tag{5.8}$$

Donde X e Y representan la distancia al eje inferior izquierdo de la pista, XX e YY representan la distancia al jugador que posee el balón, V_x y V_y representan la velocidad, H indica el jugador que posee el balón en ese instante y L es el invariante laplaciano.

6

Sistema de Reconocimiento de Actividades de Grupo

El objetivo final de los sistemas de reconocimiento de actividades de grupo consiste en disponer de un sistema entrenado, capaz de clasificar una secuencia nueva en alguna de las clases (comportamientos de grupo) disponibles en un conjunto dado.

Para ello, nuestro sistema parte de una serie de características básicas obtenidas con la ayuda de sistemas de tracking de vídeo y un experto en el dominio tratado. A partir de estas características básicas se desarrolla un proceso llamada "Explosión de características" descrito en el próximo apartado 6.1, en el que mediante la ayuda de un experto, se obtienen toda una serie de características derivadas de las básicas de partida. Estas características derivadas pueden tener un grado variable de dependencia del dominio, que a su vez estará relacionada con la dependencia al experto empleado. De esta forma, existen características derivadas útiles en la mayoría de dominios (baja dependencia del dominio) como la velocidad, la aceleración, etc. En el otro extremo tendríamos otras características concretas del dominio tratado, como en el caso del baloncesto la posición de los jugadores relativa al jugador que lleva el balón, que tendrían una alta dependencia del dominio y para la que habría que contar con la información del experto para encontrarlas.

Una vez disponemos de las características completas (tanto las básicas como las derivadas), procederemos al etiquetado de las secuencias en función de la clase a la que perteneces, esto es, el comportamiento de grupo llevado a cabo en cada una de ellas. De nuevo, en este punto es importante contar con la colaboración de un experto en el dominio en cuestión.

Tras haber procedido a la clasificación de las secuencias, debemos realizar una criba en el

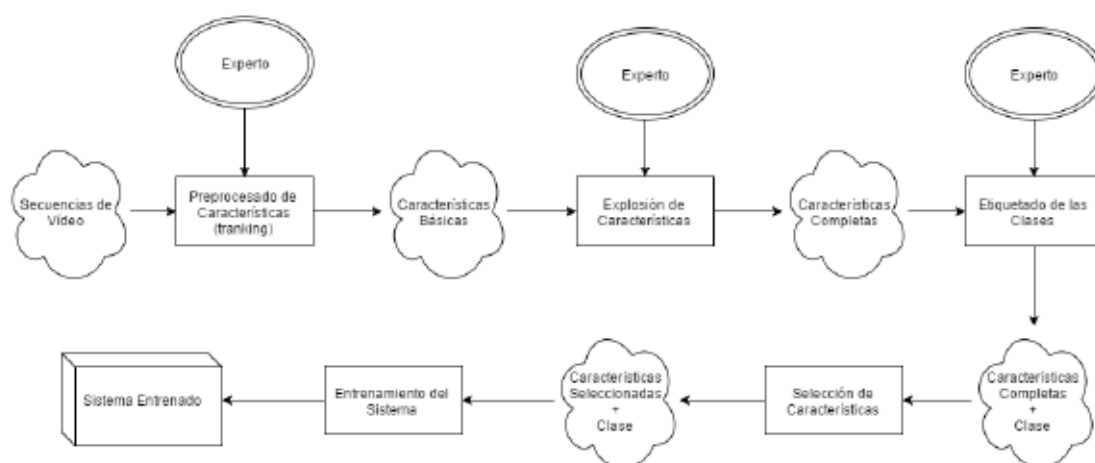


Figura 6.1: Flujo principal del entrenamiento del sistema

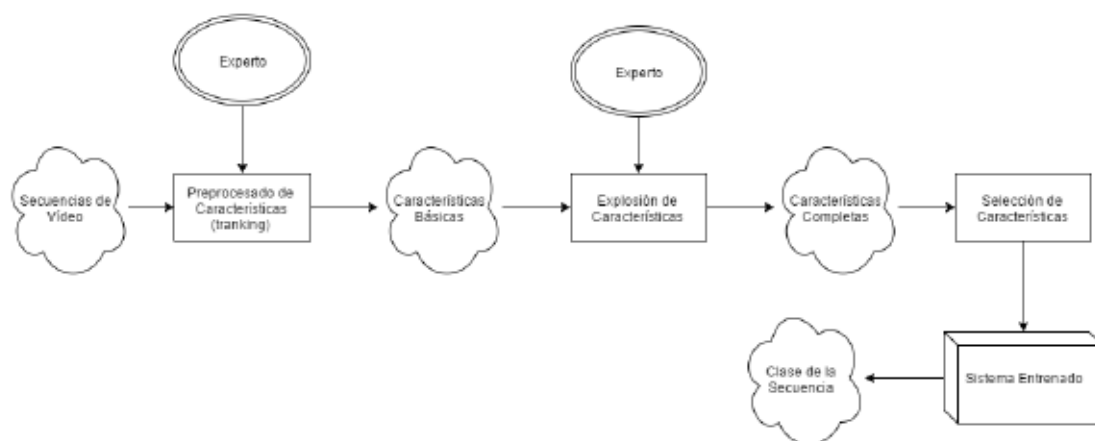


Figura 6.2: Flujo principal de la clasificación

número de las características a utilizar. Esta fase tienen una doble importancia. Por un lado, reducir el número de características reduce la dimensionalidad del problema y por lo tanto su complejidad. Por otro lado, ciertas características pueden no aportar información útil al reconocimiento de la actividad de grupo llevada a cabo sino ruido, lo que empeoraría la calidad del sistema. Obviamente, qué características se filtran y cuáles se potencian es una decisión crucial para la calidad del sistema resultante, por lo que se estudiarán diversos modos de llevar a cabo dicha selección, descritos en la sección 6.2.

Una vez que se dispone de las secuencias clasificadas según el comportamiento de grupo llevado a cabo y sus características (explotadas y filtradas) se procede a entrenar el sistema. Para ello se dispone de un Modelo Oculto de Markov (HMM) al que se le servirán los datos anteriormente relatados para su entrenamiento, una vez entrenado el sistema, se dispone de una herramienta válida para clasificar nuevas secuencias de características no introducidas en la fase de entrenamiento. Todo el proceso queda resumido en la imagen 6.1.

Tras el proceso de entrenamiento del sistema, puede procederse a clasificar nuevas secuencias y predecir así el comportamiento de grupo llevado a cabo. Para ello se repite el proceso de extracción de características, explosión y selección. Obviamente, en este caso la selección de características se lleva a cabo según los resultados obtenidos en el entrenamiento del sistema, esto es, filtrando y potenciando las mismas características que durante el proceso de entrenamiento del sistema. Una vez disponemos de las características seleccionadas podemos utilizar el sistema para clasificar la nueva secuencia, obteniendo así la clase (comportamiento de grupo) a la que pertenece. El proceso se muestra en la imagen 6.2.

6.1 Explosión de Características

En este apartado se trata de describir en detalle el proceso por el cual se pretende pasar de un subconjunto básico de características obtenidas directamente de los datos de entrada, a un conjunto de características más amplio que proporcionen una mayor cantidad de información a partir de la cual poder desarrollar todo el sistema.

Es importante tener una visión global de este proceso como un paso más dentro del sistema descrito en 6.1. una vez generado el conjunto de características resultantes del proceso de explosión de características, estas son sometidas a un proceso de filtrado, por este motivo debe tenerse más énfasis en la generación de características sin recaer demasiado en la preocupación

de que alguna de ellas acabe no siendo de utilidad, o incluso, acabe aportando ruido al sistema en lugar de información. Por ese motivo, en este proceso se pretende incorporar todas las características que, de alguna manera, se consideren que pueden ser útiles para el cometido de la clasificación, aunque no se tenga certeza de ello.

Lógicamente, este proceso tiene por una parte una cierta dependencia del dominio en cuestión, y por ello es aconsejable contar con un experto que pueda discernir qué características conviene agregar en función del dominio concreto. Sin embargo, dado que todos los contextos considerados en este estudio disponen de una serie de características comunes, puede obtenerse a raíz de esas características un conjunto básico que puede ser introducido en todas las situaciones, con independencia del dominio en cuestión.

Todos los dominios tratados parten de la información de la posición en cada instante de un conjunto de elementos móviles que conforman grupos, cuya unión fundamental es poseer un objetivo común. Por ello, características derivadas de la posición son vitales y pueden formar parte del proceso de explosión de características con independencia del dominio. Estas características serían la velocidad y la aceleración, tanto de los individuos que componen los grupos así como de los propios grupos.

Otro de los elementos comunes en todos los dominios consiste en la importancia de la "forma" de los grafos que dibujan los elementos móviles del sistema. Así, cualquier característica que ayude a definir y estudiar dicha forma puede aportar información valiosa para el sistema. Una de esas características pueden ser las componentes del invariante laplaciano, que, de forma sencilla, pueden dar una información muy valiosa sobre la forma del grafo en cada momento.

Otra característica que, si bien no está presente en todos los dominios, sí que lo está en la mayoría de ellos, es que los diferentes individuos puedan tener asignado un rol concreto dentro del grupo. Si esto es así para el dominio tratado, existen una serie de características que pueden explotarse. En este caso podemos extraer las posiciones (y velocidades y aceleraciones) de cada individuo relativas a uno en concreto. Si se dispone de información concreta del dominio puede desarrollarse estas características para un elemento concreto (seleccionado por el experto) si no, puede realizarse esta maniobra para cada uno de los roles y confiar en la selección de características del paso siguiente.

Todas estas características y aquellas que puedan obtenerse de más para un dominio concreto, son añadidas y formarán la entrada para el apartado siguiente, en el que se tratará de filtrar todas las que no aporten información valiosa, así como potenciar aquellas que puedan

ser vitales para la ejecución de la clasificación del comportamiento de grupo.

6.2 Selección de Características

Tal y como se ha comentado, el reconocimiento de actividades de grupo es usualmente un problema de una gran dimensionalidad. Por ejemplo, en el caso descrito en el apartado 5.2 se observa que de una composición de cuatro individuos y dos grupos, se han obtenido un total de cuarenta características. En caso de que se hubiera seguido el mismo procedimiento para una escena de diez individuos (un partido de baloncesto real), se hubieran obtenido un total de 76 características. Por lo tanto, puede resultar muy importante realizar una buena selección de características, que reduzca la dimensionalidad del problema de clasificación mejorando su eficacia. Además, es importante destacar que en muchas ocasiones puede ocurrir que no todas las características aporten información valiosa al sistema. Por el contrario, es muy habitual que un subconjunto de ellas pueda resultar más representativo de la clase a la que pertenece la muestra que el conjunto total de las características. Esto se debe a que en algunos casos, ciertas características pueden no aportar ninguna información valiosa sino simplemente ruido. Por este motivo, se proponen cuatro maneras diferentes de realizar la selección de las características en el reconocimiento de actividades de grupo: "wrapper global", "wrapper binario", "filtro mRMR global" y "filtro mRMR binario".

6.2.1 Wrapper global

En esta primera aproximación a la selección de características se trabaja con un único clasificador (capaz de elegir entre todas las posibles clases presentes en el conjunto de datos). Basándose en el algoritmo descrito en (Patricio and Maravall, 2000), se considera una lista inicial de características candidatas de entre las que se selecciona una única y se calcula la precisión del sistema, y se repite este proceso para cada una de las características de la lista de candidatas. La característica que obtiene mejores resultados es introducida en el conjunto de características seleccionadas, y eliminada del conjunto de candidatas. El proceso se repite de forma iterativa, realizando siempre las pruebas con la unión del conjunto de características seleccionadas y la característica del conjunto de candidatas que está siendo evaluada en cada momento. Se detiene el proceso en el momento en que la adición de ninguna característica del grupo de candidatas puede mejorar la precisión obtenida por el conjunto de características seleccionadas, o cuando

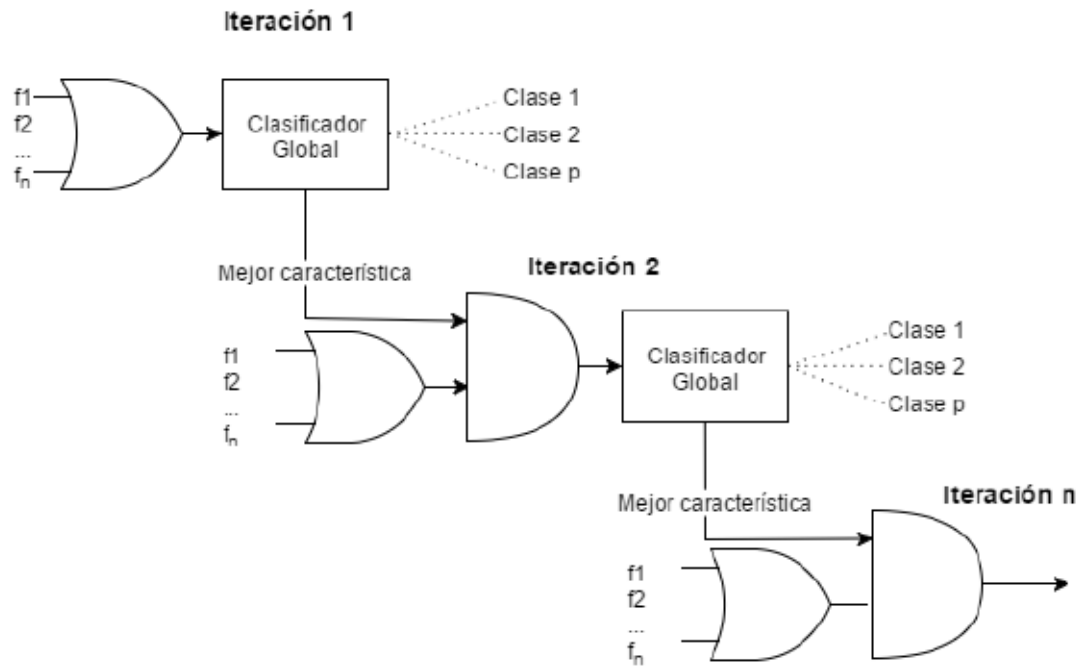


Figura 6.3: Selección de características mediante wrapper global

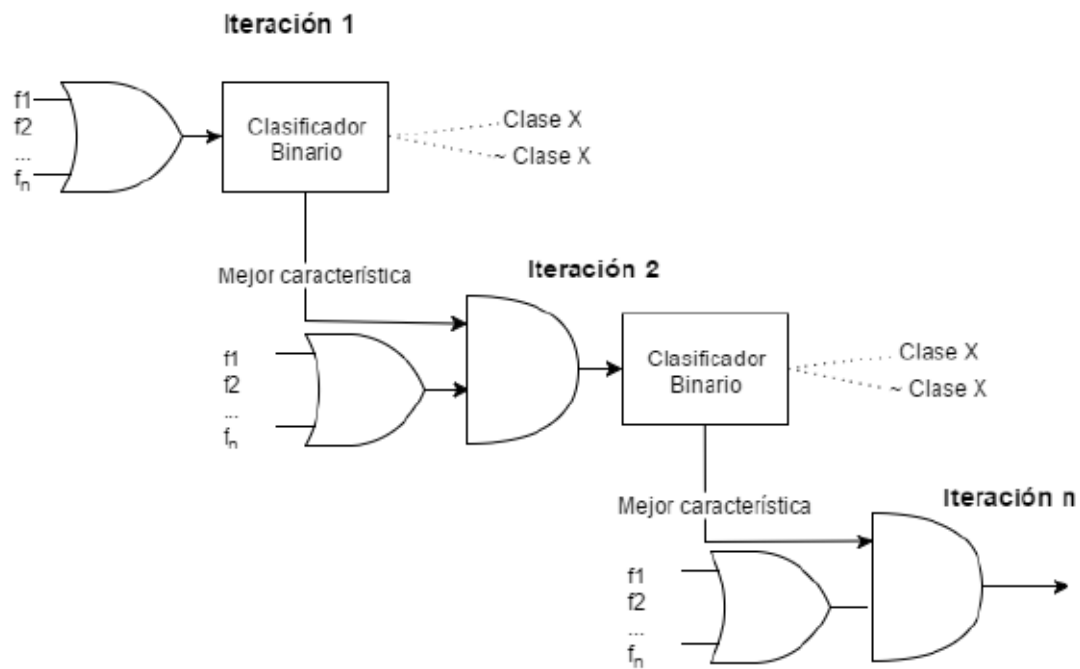


Figura 6.4: Selección de características mediante wrapper binario

se vacía la lista de características candidatas. El proceso descrito puede verse en la imagen 6.3, las puertas lógicas de tipo OR representan que sólo una característica es introducida en cada paso, mientras que las de tipo AND indican que se combinan todas.

6.2.2 Wrapper binario

Esta segunda aproximación parte de la idea descrita en el apartado anterior pero utilizando para la evaluación un conjunto de P clasificadores binarios (uno por cada clase presente en el conjunto de datos) capaz de discernir entre si una muestra pertenece o no a dicha clase. Para ello es necesario replicar el conjunto de datos tantas veces como clases existan (P), añadiendo en cada caso la etiqueta de si pertenece o no a la clase en cuestión. De esta forma cada instancia es valorada por el conjunto de todos los clasificadores, obteniendo así la probabilidad de que dicha instancia pertenezca a cada una de las clases. Finalmente, de entre los clasificadores que predijeron que la instancia sí pertenecía a la clase, se selecciona el que obtuvo mayor probabilidad. Siguiendo este sistema se realiza la selección de las mejores características para cada clasificador, siguiendo el esquema descrito en 6.2.1. Con la diferencia de que en este caso el proceso se repite por cada tipo de clase presente en el conjunto de datos, una para cada clasificador. De esta manera se dispondrá de P conjuntos de características candidatas (inicialmente similares), y de P conjuntos de características seleccionadas, cada uno de los cuales debe usarse en uno de los clasificadores binarios. La imagen 6.4 ilustra el algoritmo descrito.

6.2.3 Filtro mRMR global

En esta ocasión se ha optado por la utilización de un filtro en lugar de un wrapper para la selección de características. La diferencia fundamental entre un filtro y un wrapper, es que mientras el wrapper realiza la clasificación y utiliza la precisión obtenida para realizar la selección, un filtro trata de medir algún parámetro que discrimine qué característica es más apropiada sin tener que realizar la clasificación con cada una de ellas. En este caso, el parámetro que se ha utilizado para realizar la selección ha sido el obtenido por el algoritmo de mínima redundancia y máxima relevancia (mRMR) descrito por la ecuación 8.2, obtenido del artículo (Peng et al., 2005).

$$I(C_i, f_i) - \frac{1}{|F|} \sum_{f_s \in S} I(f_s, f_i) \quad (6.1)$$

Mediante el filtro mRMR podemos realizar una selección de características de manera mucho más rápida que con los wrapper descritos anteriormente, ya que el cálculo de la ecuación 8.2 es más rápido que el proceso de clasificación. Para el filtro mRMR global, se ha contado con un solo clasificador capaz de decidir la clase de cualquier instancia, obteniendo así un único conjunto final de características.

6.2.4 Filtro mRMR binario

En esta ocasión, se ha decidido combinar la técnica de P clasificadores descrita en el apartado 6.2.2, combinándolo con el filtro mRMR definido en la ecuación 8.2. Por lo tanto, este sistema obtiene un total de P conjuntos de características seleccionadas, una para cada clasificador, estando cada clasificador centrado en decidir si una instancia pertenece o no a una determinada clase.

6.3 Reconocimiento de la Actividad de Grupo

Una vez explicados los procesos de modelado y explosión y selección de características, en este apartado se detallan los clasificadores utilizados en el sistema. Para ello es necesario catalogar los sistemas desarrollados en dos grupos diferentes, los clasificadores globales y los clasificadores binarios. Tal y como se ha mostrado en la sección anterior, de los cuatro sistemas implementados, dos realizan clasificaciones binarias (de pertenencia o no a una determinada clase) y dos realizan clasificaciones globales (elegir de entre todas las clases aquella a la que pertenece la instancia). Ambos casos, tanto las clasificaciones globales como las binarias, pueden ser abordadas utilizando distintas técnicas de clasificación, en este caso, y dadas las características del problema, se ha optado por utilizar un modelo oculto de Markov (HMM) por su idoneidad en el reconocimiento de patrones temporales. Todo HMM viene determinado por la tupla de valores (Q, V, π, A, B) donde:

- Q representa el conjunto de estados ocultos.
- V el conjunto de los posibles valores observables.
- π las probabilidades iniciales de cada estado.
- A las probabilidades de cada una de las transiciones entre estados.
- B las probabilidades de las observaciones en función del estado actual.

Para la realización de la clasificación global (en los casos descritos en 6.2.1 y 6.2.3) se dispone de un conjunto de características (cuya dimensionalidad depende de la fase de desarrollo del sistema, pero es fija para un determinado estadio) que deben ser clasificadas dentro de una de las cinco posibles clases existentes en el dominio. Para ello se ha diseñado un HMM con tres estados ocultos por cada clase, es decir $Q = [a_1, a_2, a_3, b_1, b_2, b_3, c_1, c_2, c_3, d_1, d_2, d_3, e_1, e_2, e_3]$, V es el espacio de valores que pueden tomar las diferentes características analizadas, y π, A y B son los parámetros del sistema que deben ser entrenados mediante el algoritmo Baum-Welch (Welch, 2003). El proceso queda reflejado en la figura 6.5.

Una vez entrenado el HMM, este puede clasificar nuevas instancias de datos asignando una determinada clase, cada una de las cuales corresponde con un comportamiento de grupo determinado, mediante el algoritmo de Viterbi (Viterbi, 1967). Este proceso de entrenamiento

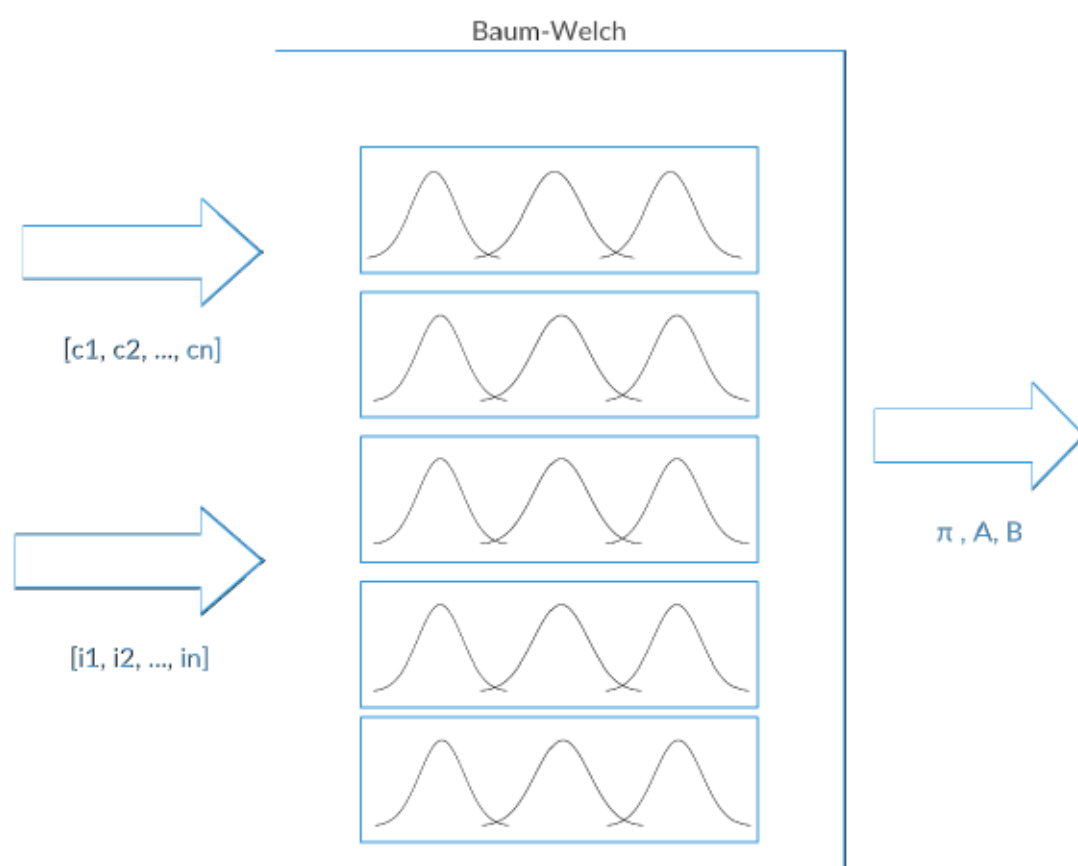


Figura 6.5: Entrenamiento del HMM global

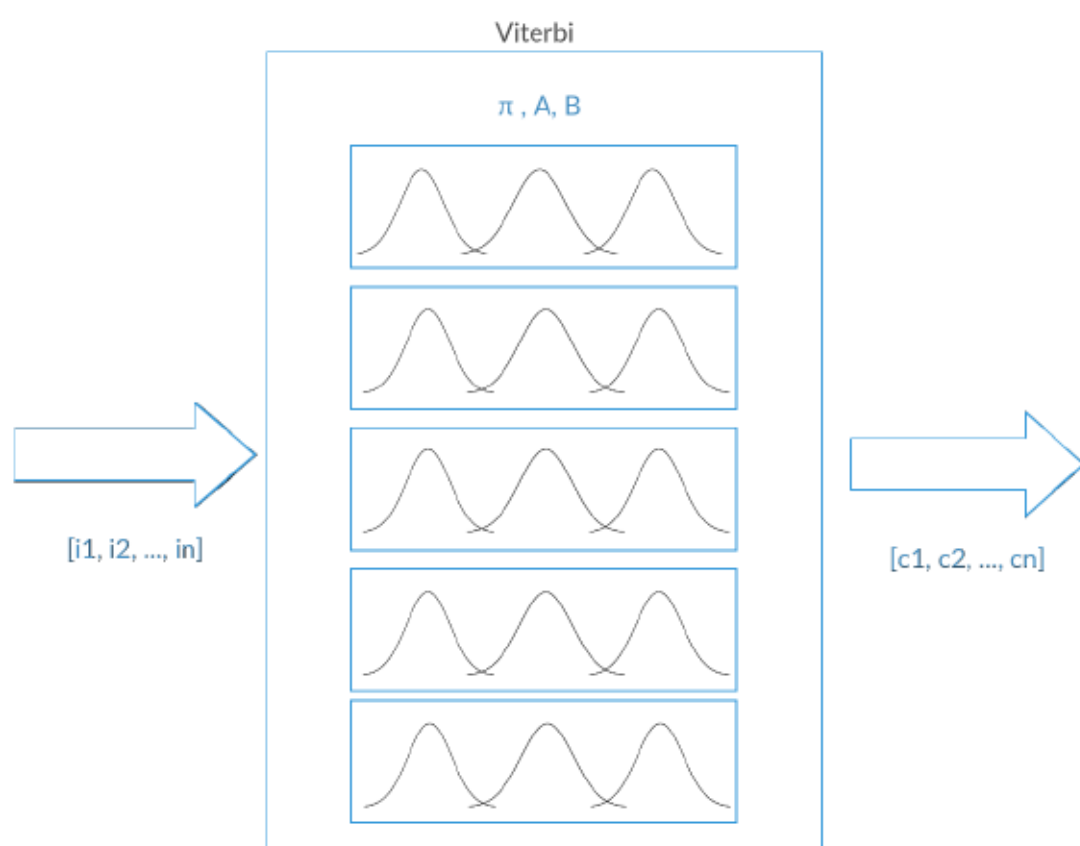


Figura 6.6: Clasificación con el HMM global

y clasificación es llevado a cabo de forma repetida, siguiendo el procedimiento descrito en 8.1.1 y en 8.1.3.

Para la clasificación de los modelos binarios (descritos en 6.2.2 y 6.2.4) se dispone de nuevo de un conjunto de características (cuya dimensionalidad depende de la fase de desarrollo del sistema, pero es fija para un determinado estadio) que, en este caso, debe ser clasificado o no dentro de una determinada clase, es decir, cada HMM es el encargado de decidir si una determinada instancia pertenece o no a una determinada clase. Para ello se ha diseñado un HMM por cada clase, con dos estados ocultos para la pertenencia a la clase y cuatro para la no pertenencia ($Q = [p_1, p_2, np_1, np_2, np_3, np_4]$). Para el entrenamiento de cada HMM se debe modificar la clase de cada instancia de entrenamiento, indicando (en vez de la clase a la que pertenece) si pertenece o no a la clase de la que se encarga ese HMM concreto. La imagen 6.7 muestra el proceso de entrenamiento.

Una vez entrenado cada uno de los HMM por separado, se juntan todos en un sistema de clasificación que obtiene las probabilidades de pertenencia a cada una de las clases para cada una de las instancias de test. De ese modo, se comparan y se clasifica la instancia con la clase que ha obtenido mayor probabilidad. La imagen 6.8 muestra un esquema del proceso de clasificación.

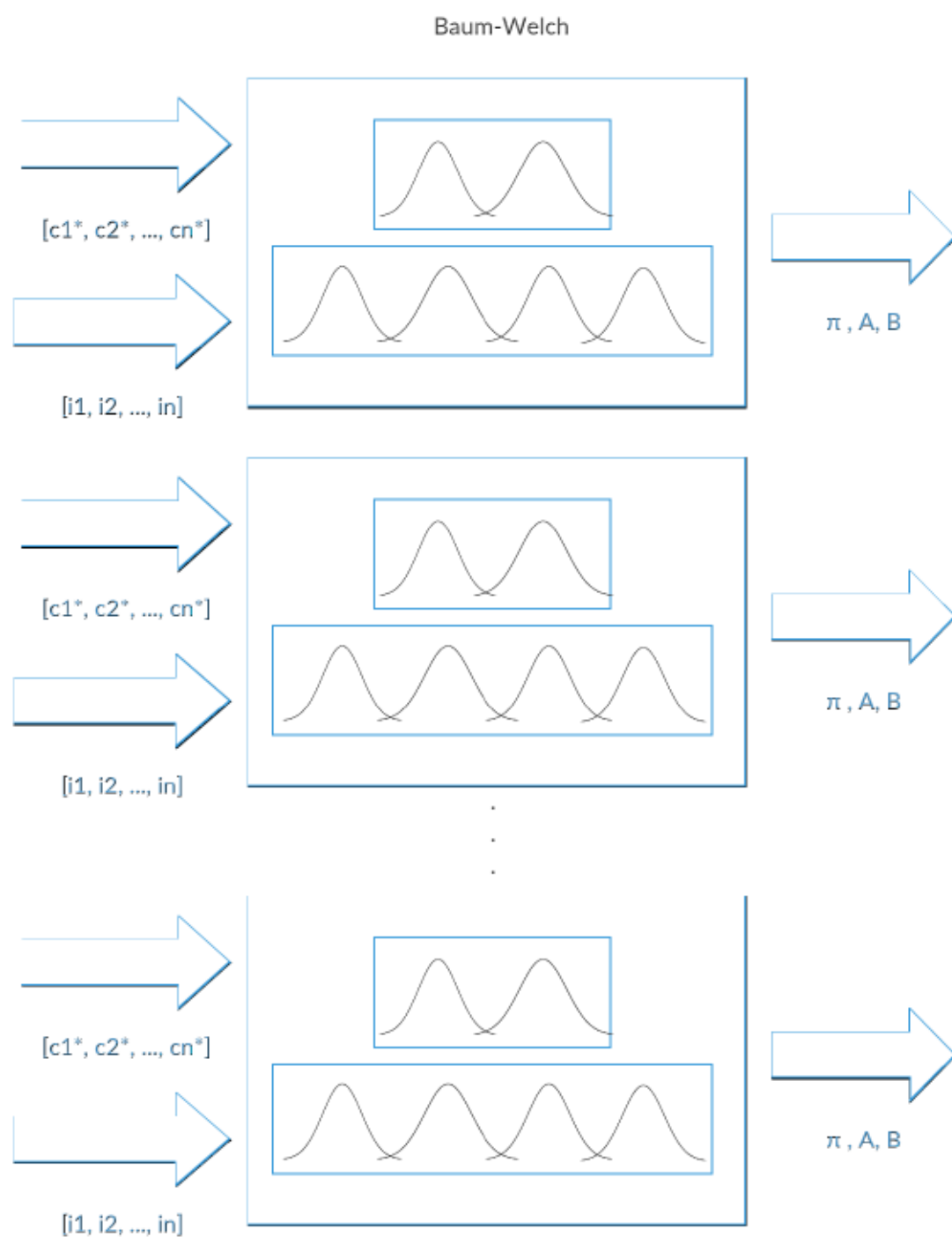


Figura 6.7: Entrenamiento del HMM binario

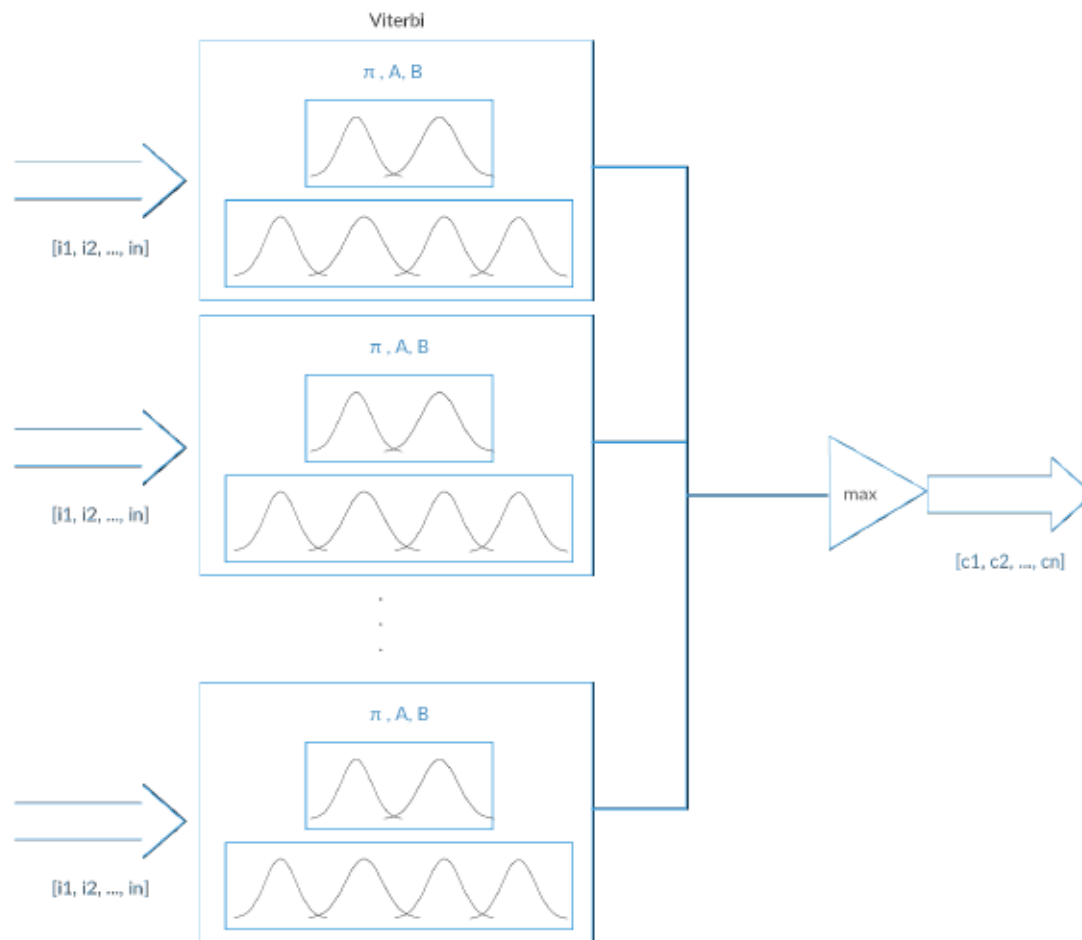


Figura 6.8: Clasificación con el HMM binario

7

Inef12Basketball Dataset

Dentro del ámbito de la inteligencia artificial, y más concretamente en el campo del reconocimiento de actividades, existen multitud de trabajos que desde diferentes puntos de vista tratan de gestar sistemas autónomos que permitan clasificar, detectar o simular el comportamiento de individuos. Todos estos trabajos basan su experimentación en un amplio conjunto de datasets, que conforman un extenso campo de investigación, y permiten a esta rama de la ciencia avanzar con cierta velocidad. Sin embargo, dentro de este campo de conocimiento existen una variación algo más novedosa, que surge como una ampliación natural del espectro de posibles problemas a tratar, y que sin embargo abre un campo muy novedoso y plagado de nuevos desafíos, alejados de alguna manera del problema inicial: el reconocimiento de actividades en grupo. En este nuevo campo el comportamiento que pretende ser clasificado, detectado o simulado no pertenece a un individuo sino a un conjunto de individuos, entre los que se establecen una serie de relaciones de diferente índole y en el que se aumenta de manera muy notable la dificultad y complejidad del problema. Esta nueva rama cuenta a día de hoy de un número todavía reducido de investigaciones en comparación a su predecesora, y esto produce que también sea escaso el número de datasets disponibles para los científicos, lo cual frena en gran medida el avance de las investigaciones. INEF12Basketball dataset (Pozo et al., 2012a) pretende contribuir con esta rama de la ciencia, aportando un conjunto de datos de bajo nivel tales como la posición de los elementos que intervienen en el sistema, así como otros de más alto nivel relacionados con el comportamiento subyacente del grupo. De esta manera se trata de conseguir un trabajo que pueda ser utilizado por los diferentes investigadores para desarrollar, mejorar y certificar sus sistemas, dentro de la rama del reconocimiento de actividades en grupo. INEF12Basketball dataset contiene una grabación de un entrenamiento de baloncesto y se ha dotado al vídeo de cierta información de alto nivel que puede ser utilizada en diferentes entornos de investigación

tales como: el reconocimiento de actividades individuales, reconocimiento de actividades en grupo, visión por ordenador, etc. Además de los vídeos y la información adicional suministrada, se incluye un conjunto de funciones Matlab que pretender servir de ayuda para la utilización y manipulación de los datos proporcionados.

7.1 Contexto

El vídeo tomado corresponde a un entrenamiento de baloncesto 2vs2, en el que se practican diferentes defensas ante un ataque de tipo pick and roll. Para ello se han dispuesto un total de veintisiete jugadas, en las que se practican ocho tipos de defensas diferentes. La estrategia de ataque pick and roll consiste en que el jugador atacante que posee el balón se aproxima a un compañero, este a su vez se queda estático en una posición tal que dificulta el avance del defensor (pantalla o pick) que cubre al portador del balón. Una vez que el jugador que dispone el balón adquiere ventaja ante su defensor, su compañero abandona la posición de pantalla (o pick) y realiza el desmarque (también llamado roll). Ante este tipo de ataque existen diferentes formas de defensa, en función del comportamiento que siga tanto el jugador que defiende al que posee el balón como el que defiende a su compañero. A continuación se describen el conjunto de defensas practicadas en este dataset:

- **Pasar por delante:** En este tipo de defensa el jugador que marca al que lleva el balón trata de evitar el bloqueo pasando por delante del mismo, tal y como se muestra en 7.1a y 7.1b. La línea blanca representa el bloqueo del atacante, y la línea azul muestra la trayectoria del jugador defensor.
- **Pasar de tercero:** Esta defensa es similar a la anterior, salvo que en esta ocasión el jugador defensor trata de esquivar el bloque pasando por detrás, tal y como se muestra en 7.2a y 7.2b. La línea azul representa el bloqueo del atacante, y la línea blanca muestra la trayectoria del jugador defensor.
- **Push:** En esta jugada el segundo defensor (el encargado de cubrir al jugador que realiza el bloqueo) trata de impedir la continuación (o roll) de éste. De esta manera impone su cuerpo tratando de impedir el acompañamiento de la jugada, e intentando expulsar hacia fuera el bloqueo. Las imágenes 7.3a y 7.3b ilustran esta situación. La línea azul representa el bloque del jugador atacante, mientras que la línea blanca representa la oposición del segundo defensor a que se realice la continuación.



(a)



(b)

Figura 7.1: Pasar por delante

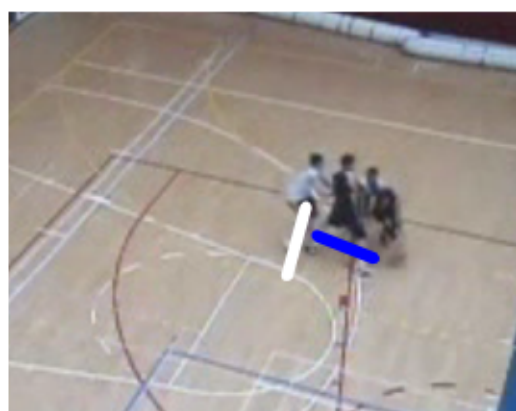


(a)

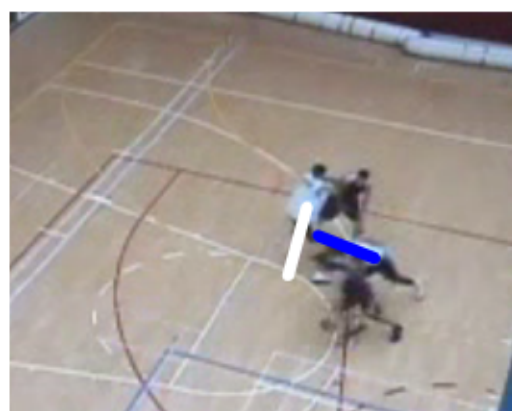


(b)

Figura 7.2: Pasar de tercero



(a)



(b)

Figura 7.3: Push



Figura 7.4: Flash lateral

- Flash lateral:** En la defensa de flash lateral el jugador encargado de cubrir al que realiza el bloqueo se desplaza lateralmente hacia la salida del jugador que conduce el balón. De esta manera proporciona una ayuda importante al defensor del balón, durante el tiempo en que este ha caído bajo el bloqueo. Una vez realizado el desplazamiento lateral el defensor vuelve con su marca inicial. Las imágenes 7.4a y 7.4b muestran el bloqueo del atacante (línea blanca) y el desplazamiento lateral del defensor (línea azul).
- Flash corto:** Durante un flash corto el defensor que cuya marca hace el bloqueo realiza una ayuda breve a su compañero, presionando la salida del jugador que tiene el balón durante un pequeño espacio de tiempo. En las imágenes 7.5a y 7.5b se muestra el bloqueo del atacante (línea blanca) y el avance del segundo defensor para presionar al jugador que conduce el balón (línea azul).
- Flash largo:** Defensa similar a la anterior en la que la ayuda se alarga más en el tiempo. De esta manera el segundo defensor sigue durante más tiempo el avance del jugador que lleva el balón, volviendo al final a su marca inicial. De nuevo en las imágenes 7.6a y 7.6b puede observarse el bloque y el avance del segundo defensor para cubrir al jugador que lleva el balón. Posteriormente se produce de nuevo el cambio entre los pares defensivos.
- Cambio:** En esta ocasión la defensa opta por realizar el cambio de los emparejamientos defensivos. Así pues el jugador que defendía al que conduce el balón queda después del bloqueo defendiendo al que ha realizado el bloqueo, cambiando la marca con su compañero. En las imágenes 7.7a y 7.7b se puede observar las flechas blancas que indican el movimiento de cada uno de los defensas siguiendo al jugador que cubría inicialmente



(a)



(b)

Figura 7.5: Flash corto



(a)



(b)

Figura 7.6: Flash largo



Figura 7.7: Cambio

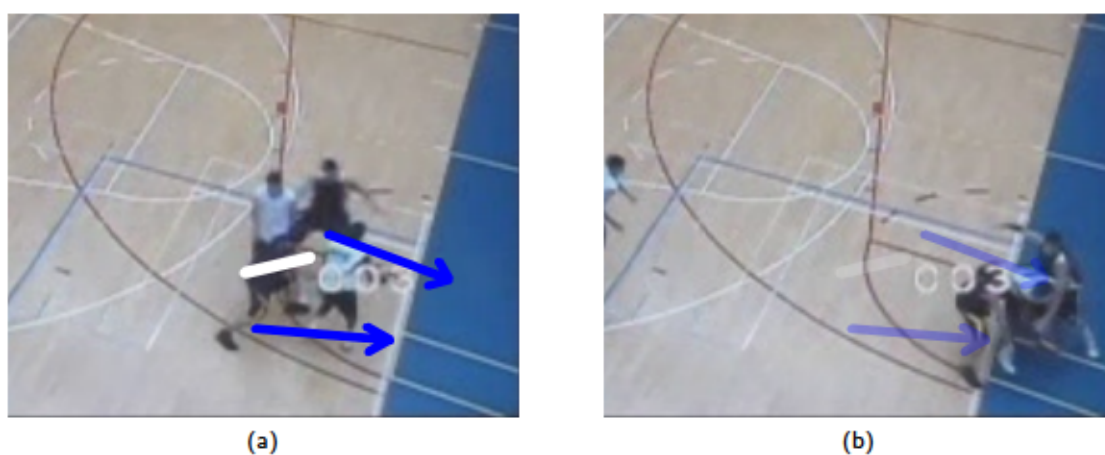


Figura 7.8: Dos contra uno

su compañero.

- **Dos contra uno:** Por último, en esta ocasión los dos defensores tras sufrir el bloqueo se quedan cubriendo al jugador que conduce el balón, tratando de obtener ventaja de una situación de dos contra uno y robar así el balón. En las imágenes 7.8a y 7.8b se muestra el bloque del atacante (línea blanca) así como el movimiento de los dos defensas que salen a buscar al jugador que tiene la posesión para provocar el robo en el dos contra uno.

Nombre	Descripción	Tipo de Datos	Ejemplo
Frames	Número de fotogramas	Entero positivo	3
Seconds	Segundo al que pertenece el fotograma	Decimal positivo	0.4
J[NUM] X	Posición del jugador NUM en el eje X según las coordenadas de la pista de baloncesto	Decimal positivo	67.7585
J[NUM] Y	Posición del jugador NUM en el eje Y según las coordenadas de la pista de baloncesto	Decimal positivo	29.3657
J[NUM] XA	Posición del jugador NUM en el eje X según las coordenadas de la imagen de la primera cámara	Decimal positivo	169.0
J[NUM] YA	Posición del jugador NUM en el eje Y según las coordenadas de la imagen de la primera cámara	Decimal positivo	192.0
J[NUM] XB	Espacio reservado para incluir una coordenada más, no presente en la actual implementación	-	-
J[NUM] YB	Espacio reservado para incluir una coordenada más, no presente en la actual implementación	-	-

Cuadro 7.1: Positions.txt

7.2 Datos

- Vídeo:** El vídeo ha sido grabado mediante cuatro cámaras situadas en posiciones fijas en una cancha de baloncesto. Existen un total de 27 jugadas grabadas, sumando un total de 3:45 minutos. A 25 fotogramas por segundo dan un total de 5.627 fotogramas. El vídeo se encuentra grabado en formato DivX Mpeg4 con contenedor avi. Sin audio asociado y con un tamaño de 320 por 240 píxeles.
- Posición de los jugadores:** La posición de cada uno de los cuatro jugadores a lo largo de los 3:45 minutos de vídeo. Esta información ha sido capturada utilizando una versión modificada del programa ViPER, mediante el uso de algoritmos de seguimiento asesorados por un experto. Para ello la posición obtenida por el algoritmo de seguimiento ha sido comprobada y supervisada por un experto, realizando las correcciones oportunas cuando era necesario. Esta información se ha almacenado en el fichero "positions.txt" dentro del directorio "basketball" del dataset. Si bien la información puede ser accedida de manera fácil y cómoda mediante las funciones de Matlab presentadas en el apartado 3, a continuación se presenta el formato en el que se encuentra recogida la información dentro del fichero "positions.txt" para facilitar así su acceso de cualquier otra forma decidida por el investigador.

En la Tabla 7.1 se pueden ver los diferentes campos de información que contiene el

Nombre	Descripción	Tipo de Datos	Ejemplo
Number	Identificador único del jugador	Entero positivo	1
Description	Descripción del rol llevado a cabo por el jugador al que pertenece el identificador, es importante destacar que los jugadores son identificados en función del papel llevado a cabo durante la jugada y no de manera personal	Texto	Attacker whose has the first ball

Cuadro 7.2: PlayerList.txt

Papel	Identificador
Jugador que empieza la jugada con el balón	1
Compañero del jugador que empieza con el balón	2
Defensor que empieza marcando al jugador 1	3
Compañero del jugador 3	4

Cuadro 7.3: Equivalencia entre identificador y papel

fichero positions.txt, así como una breve descripción, su dominio y un ejemplo ilustrativo. Es importante destacar los campos compuestos cuyo inicio es J[NUM]. En ellos se hace referencia a todos los campos cuyo nombre está compuesto por: J + número del jugador + coordenada (X o Y) + letra identificadora del eje de coordenadas. Por lo tanto, al existir en el conjunto de datos cuatro jugadores y tres posibles ejes de coordenadas, existen en un total de 24 campos. Por ejemplo, la posición en el eje x del jugador 2 según el eje de coordenadas B sería "J2 XB". Para conocer la lista de jugadores existen puede accederse al fichero "playerList.txt", cuyo formato se presenta en la siguiente tabla: Debe explicarse que la identificación de cada jugador es relativa al papel llevado a cabo en la jugada, y totalmente independiente de la persona que lo lleva a cabo. Por ejemplo el jugador 1 es el jugador que comienza la jugada con la posesión del balón, y será así aunque en cada una de las jugadas aunque la persona en concreto que lleve a cabo ese papel sea diferente. En la tabla 7.3 se muestra la equivalencia entre el papel llevado a cabo y el identificador asignado:

- **Posesión del balón:** También se dispone de la información de qué jugador tiene la posesión del balón en cada instante (en caso de tenerla alguno). Para ello, y mediante la anotación manual por parte de un experto, se almacena la información del instante de inicio y el instante de fin de cada una de las posesiones de cada uno de los jugadores dentro del fichero "hasball.txt" del directorio "basketball". La tabla 7.4 muestra el formato

Nombre	Descripción	Tipo de Datos	Ejemplo
GetBall[sec]	Instante del tiempo en el que el jugador consigue el balón, medido en segundos	Entero positivo	54
LossBall[sec]	Instante del tiempo en el que el jugador pierde el balón, medido en segundos	Entero positivo	68
Player	Número que identifica al jugador, véase "PlayerList.txt"	Entero positivo	2

Cuadro 7.4: HasBall.txt

Nombre	Descripción	Tipo de Datos	Ejemplo
Start [sec]	Instante en el que comienza la acción de grupo, medido en segundos	Entero positivo	7
End [sec]	Instante en el que termina la acción de grupo, medido en segundos	Entero positivo	14
Activity	Actividad en grupo llevada a cabo	Cualquiera de las palabras clave de actividad de grupo, véase "TeamActivityDictionary.txt"	push

Cuadro 7.5: TeamActivity.txt

con el que se ha recogido la información dentro del fichero:

- **Actividad de grupo:** Con el propósito de servir de utilidad para los campos de investigación relacionados con el reconocimiento, segmentación o clasificación de las actividades de grupo se ha incorporado la información relativa a la acción llevada a cabo por la defensa en cada instante. Así mismo, tal y como se explica en el apartado Contexto, se dispone de ocho tipos de acciones de grupo posibles, realizadas en todo caso por la pareja defensora. Esta información ha sido anotada por el especialista en baloncesto Ignacio Refoyo. Toda la información de la actividad de grupo llevada a cabo está recogida en el fichero "teamActivity.txt" cuyo formato es explicado en la tabla 7.5.

De esta manera se recogen cada una de las acciones de grupo llevadas a cabo, mediante un identificador de la acción y el intervalo de tiempo en que se realiza. La equivalencia existente entre los diferentes valores del atributo `.activityz` su significado, así como del identificador numérico utilizado en las estructuras de datos de Matlab (véase Funciones Matlab) está almacenada dentro del fichero "teamActivityDictionary.txt", cuya estructura y significado viene dado por las tablas 7.6 y 7.7 respectivamente.

Nombre	Descripción	Tipo de Datos	Ejemplo
Number	Numero de codificación de la acción grupal	Entero positivo	2
Mark	Identificador único de la acción de grupo	Texto	push
Meaning	Explicación de la acción de grupo	Texto	long flash

Cuadro 7.6: TeamActivityDictionary.txt

Mark	Acción de grupo
fp	El jugador 3 trata de evitar el bloqueo pasando por delante
thp	El jugador 3 trata de evitar el bloqueo pasando por detrás
push	El jugador 4 trata de empujar hacia fuera al jugador que realiza el bloqueo (el número 2)
sidef	El jugador 4 realiza un desplazamiento lateral procurándole una leve ayuda a su compañero en la defensa
longf	El jugador 4 realiza una ayuda en la defensa del jugador 1 de larga duración, volviendo después a su marca del jugador 2
change	Los defensas (jugadores 3 y 4) cambian su marca respectivamente
2vs1	Los defensas (jugadores 3 y 4) realizan un dos contra uno frente al jugador que tiene el balón (el jugador 1)

Cuadro 7.7: Equivalencia entre indetificador de jugada y explicación

7.3 Funciones Matlab

Todos los datos comentados en el apartado anterior se presentan acompañados de un conjunto de funciones de Matlab que facilitan el acceso a los mismos. A continuación se describen cada una de las funciones, con sus respectivos atributos de entrada y la salida que proporcionan:

- **INEF12BasketballInit:**

- Descripción: Función que se encarga de inicializar todos los datos. Debe ser llamada antes de utilizar cualquiera de las otras funciones proporcionadas en el dataset. En ella se rellenan cada una de las variables globales necesarias para el funcionamiento del resto de funciones.
- Estructura: INEF12BasketballInit (pth)
- Atributos de entrada:
 - * pht: Ruta en la que se encuentra el directorio raíz del dataset.
- Salida: Esta función no produce ninguna salida más allá de rellenar las variables globales del dataset.

- **INEF12BasketballGetPos:**

- Descripción: Función que obtiene la posición de un jugador determinado en todos y cada uno de los fotogramas del video.
- Estructura: `[X,Y,XA,YA,XB,YB, VALIDREF] = INEF12BasketballGetPos (player-num)`.
- Atributos de entrada:
 - * `playernum`: Identificador del jugador del que se quiere obtener la posición. La tabla 7.3 muestra la equivalencia entre identificador y rol del jugador.
- Salida: Devuelve la posición del jugador indicado en cada uno de los fotogramas de la imagen. La posición puede ser dada en tres sistemas de coordenadas diferentes, en el caso actual se encuentran rellenos los dos primeros, correspondientes a la posición en la cancha de juego y en la primera cámara.
 - * `X`: Vector de dimensión igual al número de fotogramas (5.627), que indica la distancia del jugador a la línea de fondo de la cancha de baloncesto.
 - * `Y`: Vector de dimensión igual al número de fotogramas (5.627), que indica la distancia del jugador a la línea lateral superior de la cancha.
 - * `XA`: Vector de dimensión igual al número de fotogramas (5.627), que indica la distancia del jugador al margen izquierdo de la imagen de la primera cámara.
 - * `YA`: Vector de dimensión igual al número de fotogramas (5.627), que indica la distancia del jugador al margen superior de la imagen de la primera cámara.
 - * `XB`: Vector reservado para almacenar la posición en un nuevo eje de coordenadas. Actualmente no soportado por los datos.
 - * `YB`: Vector reservado para almacenar la posición en un nuevo eje de coordenadas. Actualmente no soportado por los datos.

- **INEF12BasketballShowTrajectories:**

- Descripción: Función utilizada para mostrar por pantalla la trayectoria del jugador indicado en el atributo de entrada. También puede mostrarse sobreimpresionada la trayectoria de todos los jugadores del sistema.
- Estructura: `INEF12BasketballShowTrajectories (playernum)`.
- Atributos de entrada:

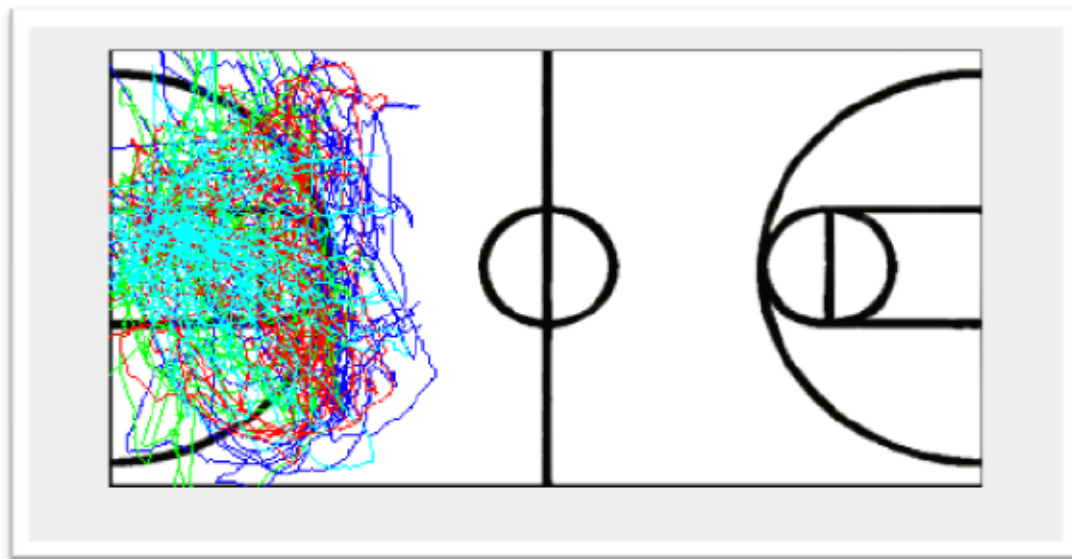


Figura 7.9: Plotter con la trayectoria de los jugadores (INEF12BasketballShowTrajectories())

- * **playernum:** Identificador del jugador del que se quiere mostrar la trayectoria. Véase la TABLA 1. En caso de no especificar ningún identificador, esta función muestra la trayectoria de todos los jugadores presentes en el sistema.
- **Salida:** Esta función no devuelve ningún valor, pero muestra por pantalla la trayectoria del jugador seleccionado, o la de todos los jugadores en caso de no seleccionar ninguno.

- **INEF12BasketballGetFrames:**

- **Descripción:** Obtiene en sendas matrices las imágenes pertenecientes al fotograma seleccionado en las distintas cámaras del sistema. Además del instante de tiempo al que pertenece.
- **Estructura:** $[A, B, C, D, T] = \text{INEF12BasketballGetFrames}(\text{framenum})$.
- **Atributos de entrada:**
 - * **framenum:** Número de fotograma que se desea obtener, en los datos proporcionados puede ir desde uno hasta 5.627.
- **Salida:**
 - * **A:** Matriz de $240 \times 320 \times 3$ que representa el fotograma indicado en el atributo de entrada según la primera cámara.

- * B : Matriz de 240x320x3 que representa el fotograma indicado en el atributo de entrada según la segunda cámara.
- * C: Matriz de 240x320x3 que representa el fotograma indicado en el atributo de entrada según la tercera cámara.
- * D: Matriz de 240x320x3 que representa el fotograma indicado en el atributo de entrada según la cuarta cámara.
- * T: Instante de tiempo (medido en segundos) al que pertenece el fotograma seleccionado.

- **INEF12BasketballShowFrames:**

- Descripción: Función empleada en mostrar por pantalla la imagen perteneciente al fotograma seleccionado en cada una de las diferentes cámaras. Además en la primera cámara aparece situada la posición de cada uno de los jugadores.
- Estructura: INEF12BasketballShowFrames (framenum)
- Atributos de entrada: framenum: Número de fotograma que se desea mostrar.
- Salida: Esta función no devuelve ningún resultado. Muestra por pantalla la imagen correspondiente al fotograma seleccionado en cada una de las cámaras del sistema. Además de la posición de los jugadores situada en la primera de las cámaras. A continuación se muestra un ejemplo de ejecución:

- **INEF12BasketballGetTeamActivity:**

- Descripción: Función empleada para obtener los datos de la actividad de grupo realizada en cada momento.
- Estructura: [Timestart, Timeend, Framestart, Frameend, Act] = INEF12BasketballGetTeamActivity.
- Atributos de entrada: No cuenta con atributos de entrada
- Salida:
 - * Timestart: Vector que contiene el instante (en segundos) en que comienza cada una de las acciones de grupo.
 - * Timeend: Vector que contiene el instante (en segundos) en que finaliza cada una de las acciones de grupo.

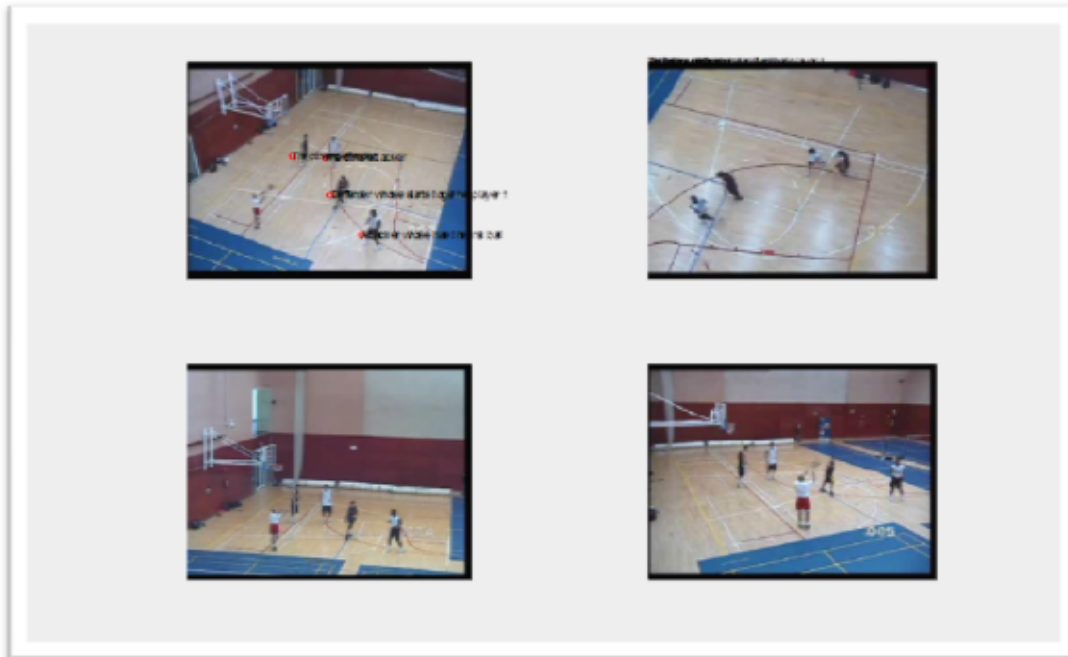


Figura 7.10: Ejemplo de ejecución de la función ShowFrames

- * **FRAMESTART**: Vector que contiene el número de fotograma en que comienza cada una de las acciones de grupo.
- * **TIMEEND**: Vector que contiene el número de fotograma en que finaliza cada una de las acciones de grupo.
- * **ACT**: Vector que contiene cada una de las acciones llevadas a cabo por el grupo. Las acciones grupales son identificadas por un número entero positivo, la equivalencia entre el número de identificador y la acción grupal puede comprobarse en el fichero "TeamActivityDictionary.txt".

- **INEF12BasketballHasBall:**

- Descripción: Función utilizada para obtener la información sobre qué jugador tiene la posesión del balón en cada momento. Esta es, a excepción de la posición, la única información individual que se incorpora en el conjunto de datos.
- Estructura: [TIMESTART,TIMEEND,FRAMESTART,FRAMEEND] =
= INEF12BasketballHasBall (playernum).
- Atributos de entrada:
 - * playernum: Identificador del jugador del que se quiere conocer en qué intervalos

a poseído el balón. Véase 7.1.

– Salida:

- * **TIMESTART**: Vector que contiene los instantes de tiempo (medidos en segundos) en los que el jugador consultado adquiere el balón.
- * **TIMEEND**: Vector que contiene los instantes de tiempo (medidos en segundos) en los que el jugador pierde la posesión del balón.
- * **FRAMESTART**: Vector que contiene el número de los fotogramas en los que el jugador consultado adquiere el balón.
- * **TIMESTART**: Vector que contiene el número de los fotogramas en los que el jugador consultado adquiere el balón.



Resultados Obtenidos

A fin de validar la propuesta realizada en el apartado 6, se han realizado una serie de experimentos utilizando el conjunto de datos descrito en 7 así como los protocolos y métricas descritos en el apéndice B. A continuación se detallan dichos experimentos, así como los resultados obtenidos.

8

Sistema de Reconocimiento de Actividades de Grupo

La propuesta llevada a cabo en este anteproyecto de tesis será evaluada empleando el conjunto de datos *INEF12Basketball Dataset*.³¹ que se puede acceder de forma pública en la dirección web: <https://sourceforge.net/projects/inef12dataset/>.

El protocolo de evaluación seguido para comprobar el rendimiento de los métodos propuestos será el de validación cruzada dejando uno fuera o "Leave One Out - Cross-Validation". Los modelos serán entrenados dejando siempre una instancia fuera, la cual será empleada en la validación.

En dichos experimentos se tomará como medida de rendimiento la tasa de reconocimiento obtenida, siendo esta el número de instancias correctamente clasificadas partido del número de instancias totales, tal y como se muestra en la ecuación B.1.

$$TasadeReconocimiento = \frac{\# \text{ de predicciones correctas}}{\# \text{ total de instancias}} \quad (8.1)$$

Esta medida puede considerarse estándar en el campo del reconocimiento de actividades en grupo, y resultar así de método valioso para la comparación con otros trabajos y aproximaciones.

8.1 Resultados Obtenidos

Una de las dificultades presentes en la actualidad en el reconocimiento de actividades de grupo es la escasez de conjuntos de datos válidos para entrenar y testear los sistemas. Por ese motivo, para el desarrollo de esta tesis doctoral se ha creado un conjunto de datos específico tal y como se ha expuesto en el apartado 7.

Todos experimentos han sido desarrollados en Matlab. Todos los algoritmos descritos han sido implementados en Matlab, por su idoneidad para esta tarea y porque el conjunto de datos utilizado 7 provee en dicho lenguaje de todo lo necesario para ser utilizado.

En primer lugar se han cargado todos los elementos del conjunto de datos, el cual proporciona un total de 27 instancias cada una de las cuales pertenecientes a un total de ocho comportamientos de grupo diferentes. Tras analizar los datos, se ha concluido que debían eliminarse tres de esos comportamientos puesto que no disponían de un número de repeticiones suficiente para los algoritmos de aprendizaje automático. Tras esta selección inicial, han resultado un total de 23 instancias pertenecientes a cinco comportamientos grupales diferentes.

Utilizando las funciones provistas en 7, cargamos los datos y disponemos de una Matriz M_i por cada instancia ($1 \leq i \leq 23$) y un vector de etiquetas indicando la clase de cada una de ellas. Cada M_i matriz tienen una dimensionalidad de $F \times T_i$, donde F es el número de características seleccionadas y T_i es la duración (en número de fotogramas) del comportamiento i . Esta matriz es la entrada del algoritmo descrito anteriormente, que será evaluado con la técnica de "Leave One Out - Cross-Validation".

Una vez disponemos de todos los elementos necesarios dentro del framework de Matlab, podemos proceder a realizar el proceso de explosión de características descrito en 6.1 obteniendo el conjunto de cuarenta características mostradas en la tabla 5.1.

Para la clasificación se ha dispuesto de la técnica Hidden Markov Model (HMM), con un número de gaussianas por clase que depende del tipo de filtro de características utilizado (véase 6.2). A continuación se muestran todos los detalles de la experimentación para cada uno de los tipos de filtros implementados (número de gaussianas del HMM, características seleccionadas, resultados obtenidos, etc.)

		Posición jugadores	Posición equipos	Posición todos	Dueño balón	Posición jugadores (relativa)	Velocidad Jugadores	Velocidad equipos	Velocidad todos	Invariante Laplaciano																														
Pasar delante																																								
Pasar detrás																																								
Push																																								
Flash largo																																								
2 vs 1																																								
#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40

Figura 8.1: Matriz de clasificación para la primera iteración del wrapper global

8.1.1 Resultados obtenidos aplicando Wrapper global

Se dispone de un conjunto de cuarenta características a lo largo del tiempo, del que debemos seleccionar un subconjunto que permita una buena clasificación de la acción de grupo (defensa puesta en marcha) llevada a cabo en cada secuencia. Basándose en el algoritmo descrito en (Patricio and Maravall, 2000), se considera una lista inicial de características candidatas (toda la tabla 5.1). Utilizando una única característica se calcula la precisión del sistema, y se repite este proceso para cada una de las características de la lista de candidatas. La característica que obtiene mejores resultados es introducida en el conjunto de características seleccionadas, y eliminada del conjunto de candidatas. El proceso se repite de forma iterativa, realizando siempre las pruebas con la unión del conjunto de características seleccionadas y la característica del conjunto de candidatas que está siendo evaluado en cada momento, véase 6.3. Se detiene el proceso en el momento en que la adición de ninguna característica del grupo de candidatas puede mejorar la precisión obtenida por el conjunto de características seleccionadas, o cuando se vacía la lista de características candidatas. Todos los procesos de clasificación se llevan a cabo con un HMM de tres gaussianas para cada clase.

En la figura 8.1 pueden verse los resultados obtenidos en la primera iteración, cada fila representa una secuencia a clasificar, y cada columna representa la clasificación hecha con

k	Característica añadida	Precisión del sistema
1	Coordenada X del jugador 4 respecto al dueño del balón	64 %
2	Coordenada X del centro de masas del equipo defensor	68 %
3	Velocidad en el eje X del equipo defensor	70 %
4	Coordenada X del jugador 3 respecto al dueño del balón	74 %
5	Velocidad en el eje X del conjunto de todos los jugadores	78 %
6	Dueño actual del balón	78 %

Cuadro 8.1: Selección de características basadas en wrapper global

la característica indicada, el color de las celdas viene determinado por si se ha clasificado correctamente la instancia (verde) o no (rojo).

En la tabla 8.1 se muestran las características que se han ido añadiendo mediante este proceso, mostrando en la primera columna el número de iteración, y en la tercera columna la precisión obtenido por el conjunto de características presentes en esa iteración. Tal y como indica la tabla entre las iteraciones cinco y seis no se produce mejora, por lo que el sistema final quedaría con el subconjunto de características presentes en la iteración cinco.

8.1.2 Resultados obtenidos aplicando Wrapper binario

Para el wrapper binario se ha diseñado un sistema diferente, con cinco clasificadores binarios, cada uno especializado en descubrir si una instancia pertenece o no a una determinada clase. Cada clasificador se compone de un HMM con dos gaussianas para la clase de sí pertenencia y cuatro para la de no pertenencia. Para ello se ha replicado cinco veces el conjunto de entrenamiento, una por cada clase, cambiando la etiqueta de la clase de cada instancia por una que indica si pertenece o no a esa clase. Así, el clasificado de la clase b_n recibe un conjunto de datos en el que para cada instancia se indica si pertenece (1) o no (0) a la clase b_n . De esta manera cada instancia es valorada por los cinco clasificadores, obteniendo así la probabilidad de que dicha instancia pertenezca a cada una de las clases. Finalmente, de entre los clasificadores que predijeron que la instancia sí pertenecía a la clase, se selecciona el que obtuvo mayor probabilidad, véase 6.3. Siguiendo este sistema se realiza la selección de las mejoras características para cada clasificador, siguiendo el esquema descrito en 8.1.1. Con la diferencia de que en este caso el proceso se produce cinco veces, una para cada clasificador. De esta manera se dispondrá de cinco conjuntos de características candidatas (inicialmente similares), y de cinco conjuntos de características seleccionadas, cada uno de los cuales debe

Clase	k = 1	k = 2	k = 3	k = 4
Pasar por delante	5	7	33	-
Pasar por detrás	6	5	38	-
Push	1	30	22	-
Flash largo	27	26	32	-
Dos contra uno	11	23	26	33

Cuadro 8.2: Selección de características basadas en wrapper binario

Iteración	k = 1	k = 2	k = 3	k = 4
Precisión	70 %	91 %	96 %	96 %

Cuadro 8.3: Precisión obtenida en cada iteración con wrapper binario

usarse en uno de los clasificadores binarios.

Tras realizarse el proceso para cada una de las cinco clases presentes en el conjunto de datos, la tabla 8.2 muestra las características que se han ido añadiendo a cada subconjunto (uno por cada clase) y en 8.3 la precisión total obtenido por el sistema en cada iteración. Debido a cuestiones de maquetado, en vez del nombre de cada característica aparece el número con que se referencia en la tabla 5.1.

8.1.3 Resultados obtenidos aplicando filtro mRMR global

En esta ocasión se ha optado por la utilización de un filtro en lugar de un wrapper para la selección de características. La diferencia fundamental entre un filtro y un wrapper, es que mientras el wrapper realiza la clasificación y utiliza la precisión obtenida para realizar la clasificación, un filtro trata de medir algún parámetro que discrimine qué característica es más apropiada sin tener que realizar la clasificación con cada una de ellas. En este caso, el parámetro que se ha utilizado para realizar la selección ha sido el obtenido por el algoritmo de mínima redundancia y máxima relevancia (mRMR) descrito por la ecuación 8.2, obtenido del artículo (Peng et al., 2005).

$$I(C_i, f_i) - \frac{1}{|F|} \sum_{f_s \in S} I(f_s, f_i) \quad (8.2)$$

Mediante el filtro mRMR podemos realizar una selección de características de manera mucho más rápida que con los wrapper descritos anteriormente, ya que el cálculo de la ecuación 8.2

k	Característica añadida	Precisión del sistema
1	Coordenada X del jugador 1	39 %
2	Primera componente laplaciana del grafo de los jugadores	57 %
3	Segunda componente laplaciana del grafo de los jugadores	61 %
4	Tercera componente laplaciana del grafo de los jugadores	61 %
5	Dueño actual del balón	61 %
6	Velocidad en el eje Y del equipo defensor	52 %
7	Coordenada Y del jugador 4	43 %

Cuadro 8.4: Selección de características basadas en filtro mRMR global

Clase	k = 1	k = 2	k = 3	k = 4	k = 5
Pasar por delante	17	38	40	39	35
Pasar por detrás	20	38	40	39	37
Push	1	38	40	39	15
Flash largo	1	38	40	39	15
Dos contra uno	32	38	40	39	35

Cuadro 8.5: Selección de características basadas en filtro binario

es más rápido que el proceso de clasificación. Para el filtro mRMR global, se ha contado con un solo clasificador capaz de decidir la clase de cualquier instancia, conformado por un HMM con tres gaussianas para cada clase, obteniendo así la lista de características mostradas en la tabla 8.4 con su respectiva precisión. En este caso, se puede observar que la inclusión de más características a partir de la iteración tres no mejora el sistema, llegando incluso a empeorarlo.

8.1.4 Resultados obtenidos aplicando filtro mRMR binario

En esta ocasión, se ha decidido combinar la técnica de cinco clasificadores descrita en el apartado 8.1.2 (HMM de dos y cuatro gaussianas respectivamente), combinándolo con el filtro mRMR descrito en el apartado 8.1.3. Por lo tanto, este sistema obtiene un total de cinco conjuntos de características seleccionadas, una para cada clasificador, estando cada clasificador centrado en decidir si una instancia pertenece o no a una determinada clase. Es importante destacar que el sistema de cinco clasificadores binarios conlleva un tiempo mayor en el proceso de selección de características, por lo que el tiempo ganado al utilizar el filtro mRMR en lugar del wrapper puede ser importante.

La tabla 8.5 muestra las características seleccionadas para cada clase en cada iteración, de nuevo, por cuestiones de maquetación se ha indicado la característica en base a su identificador

Iteración	k = 1	k = 2	k = 3	k = 4	k = 5
Precisión	32 %	34 %	39 %	48 %	43 %

Cuadro 8.6: Precisión obtenida en cada iteración con filtro binario

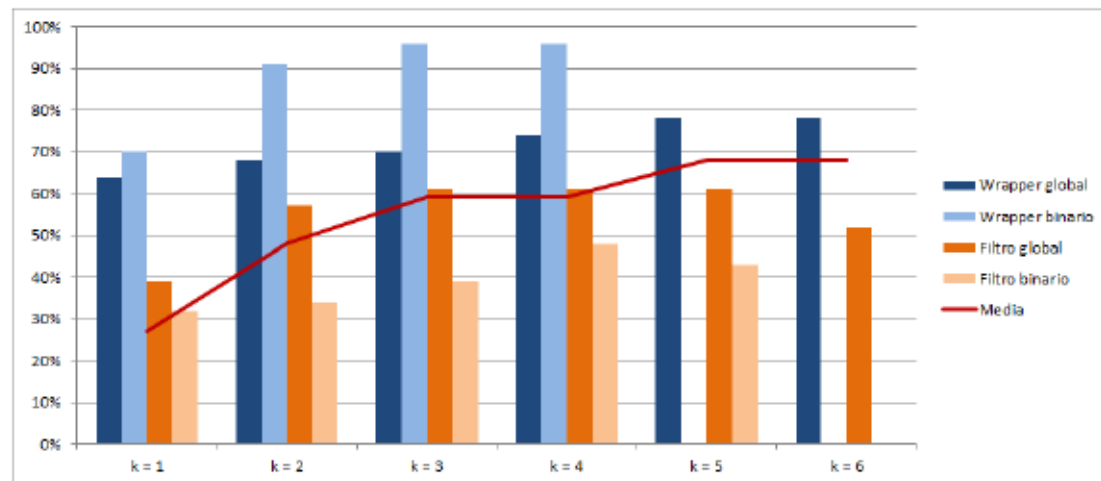


Figura 8.2: Resumen de los resultados obtenidos con los diferentes algoritmos

de la tabla 5.1. En la tabla 8.6 se muestra la precisión del sistema para cada iteración.

8.1.5 Resumen de los resultados

Una vez se han estudiado los resultados obtenidos para cada tipo de selección de características, conviene poner todos juntos para comparar y poder obtener algunas conclusiones valiosas.

La figura 8.2 muestra un resumen de los resultados obtenidos por los diferentes algoritmos en base a la dimensionalidad del problema (número de características seleccionadas), así como la media de los resultados obtenidos con todas las combinaciones posibles de ese número de características.

Tal y como se puede observar, el algoritmo basado en el wrapper binario es el que obtiene unos resultados más elevados, llegando a precisiones del 96 % para una dimensionalidad igual a cuatro y destacando por encima de los resultados medios desde el principio. Es importante tener en cuenta que este algoritmo también es el que consume más recursos, por lo que en función del problema en cuestión puede ser acertado seleccionar otro más eficiente aunque suponga una merma en los resultados obtenidos. A este respecto, cabe destacar los resultados obtenidos por la técnica del filtro global, que con tan solo una dimensionalidad igual a dos es capaz de obtener unos resultados del 61 % quedando por encima de la media. Sin embargo,

este mismo algoritmo cuando aumenta la dimensionalidad del problema empieza a quedar por debajo, por lo que parece un buen método para seleccionar características en los casos en los que se necesite un sistema rápido y con una baja dimensionalidad.

9

Conclusiones

El reconocimiento de actividades de grupo es un campo de investigación con un creciente interés, tanto científico como comercial. Cada vez son más los equipos de deportes de élite como el fútbol o el baloncesto que contratan a empresas especializadas en el tema como Amisco o Prozone para realizar análisis complejos sobre las actividades de grupo propias y de los rivales. Además existen una gran cantidad de publicaciones científicas centradas en este tema en los últimos años. Por lo tanto podemos concluir que se trata de un campo de investigación con un gran potencial, donde se deben esperar grandes avances en los próximos años.

Existen sin duda una gran variedad de enfoques diferentes, que pueden categorizarse en tres grandes áreas: Geométrica, lógica y bayesiana. Cada uno de estos enfoques ha sido analizado siguiendo los trabajos más representativos de cada uno, viendo las ventajas e inconvenientes que presentan. Esta gran variedad de enfoques se combina también con una gran variedad en las características que pueden ser obtenidas para describir el sistema, sin embargo, pese a esta gran variedad, puede concluirse que la posición de los individuos que componen los grupos es una característica principal que puede resultar fundamental en la inmensa mayoría de las situaciones.

Sin embargo, pese a la gran actividad dentro del campo de investigación, existe un claro déficit de conjuntos de datos públicos y adecuados para su uso en este área científica. Muchos de los trabajos más importantes se encuentran realizados sobre conjuntos de datos de acceso restringido, haciendo así imposible la validación o comparación con otros sistemas. Por este motivo, en la tesis propuesta en este documento se ha incorporado la creación de un conjunto de datos público y perfectamente adaptado para las necesidades del reconocimiento de actividades de grupo. En él, se ha categorizado un total de 27 jugadas distintas en el entorno del dos contra

dos de baloncesto. En el que ante el mismo ataque: "bloqueo y continuación" se desarrollan un total de cinco tipos de defensas diferentes. Además, se pone a disposición de todos los investigadores interesados un conjunto de características que sitúan a cada jugador en cada instante, además de la información de contexto de quién lleva el balón en cada instante. Este conjunto de datos público puede ser accesible para futuros trabajos de investigación relacionados con el reconocimiento de actividades de grupo, y ha servido además, para validar los sistemas presentados en esta tesis.

Debido a que la propia naturaleza del reconocimiento de actividades de grupo implica la existencia de varios individuos, y dado que de esos individuos se deben recoger un conjunto de características diferentes, nos encontramos ante un área en la que los sistemas tienen una gran dimensionalidad. Este fenómeno resulta un reto para el desarrollo de soluciones válidas. En la tesis aquí propuesta, se ha optado por realizar una selección de características para afrontar el problema de la alta dimensionalidad. Para ello, se ha descrito el proceso por el cual se pasa de un conjunto de características inicial a un conjunto mayor, con características derivadas, para en última instancia, seleccionar de entre este conjunto mayor un número reducido de ellas. Este conjunto reducido de características es el que después es utilizado para realizar el reconocimiento. Este proceso de explosión y posterior selección de las características ha sido ampliamente descrito, mostrando además cuatro variantes diferentes en el proceso de selección de las características: "Filtro Global", "Filtro Binario", "Wrapper Global", "Wrapper Binario". Además, se han presentado resultados que demuestran la eficacia de cada uno de estos sistemas, arrojando en el caso del Wrapper Binario resultados de un 96 % de acierto en la clasificación.

Por lo tanto, en el transcurso de esta tesis se ha presentado una propuesta de modelado de los problemas de reconocimiento de actividades de grupo, un sistema de explosión y posterior selección de características (con sus cuatro posibles variantes) validando la propuesta gracias a los buenos resultados obtenidos sobre el conjunto de datos Inef12Basketball, también desarrollado en esta tesis.

A

Trabajos publicados

EN este apéndice se muestra una lista con los trabajos científicos publicados a lo largo de la realización de esta tesis, organizados en función del tipo de publicación.

Artículos

- Pozo, A., Patricio, M. A., García, J., and Molina, J. M. (2013). A feature selection approach to the group behavior recognition issue using static context information. *International Journal of Distributed Sensor Networks*, 2013
- Pozo, A., García, J., and Patricio, M. (2012a). Inef12basketball dataset and the group behavior recognition issue. *Distributed Computing and Artificial Intelligence*, pages 151–160
- Pozo, A., Patricio Guisado, M. Á., García, J., Molina, J. M., and Refoyo, I. (2012b). Group behavior recognition issue, feature analysis on defending pick and roll basketball move

Publicaciones en conferencias

- Pozo, A., Gracía, J., Patricio, M. A., and Molina, J. M. (2011b). A structured representation to the group behavior recognition issue. In *User-Centric Technologies and Applications*, pages 47–57. Springer
- Pozo, A., Gracía, J., Patricio, M. A., and Molina, J. M. (2011a). Group behavior recognition in context-aware systems. In *Advances in Computational Intelligence*, pages 645–652. Springer

B

Protocolos y Métodos de Evaluación

EN este apéndice se detallan los protocolos de evaluación empleados para la obtención de resultados sobre la propuesta presentada en la tesis, así como la métrica que deben ser empleadas para la comparación entre diferentes sistemas.

B.1 Validación cruzada dejando uno fuera

La validación cruzada dejando uno fuera (en inglés "Leave-one-out cross-validation o "LOOCV") consiste en separar los datos de tal manera que en cada iteración se disponga de una única muestra de prueba y del resto de datos para el conjunto de entrenamiento. Este proceso se repite hasta que todos las instancias de los datos han sido utilizadas una vez para el conjunto de prueba.

B.2 Métricas

La métrica seleccionada para medir la precisión del sistema consiste en la tasa de reconocimiento obtenida, es decir, la proporción de instancias de prueba bien clasificadas en relación al número total de instancias. Esto es:

$$TasadeReconocimiento = \frac{\# \text{ de predicciones correctas}}{\# \text{ total de instancias}} \quad (B.1)$$

Esta métrica es comúnmente aceptada y utilizada por muy diversos trabajos por lo que resulta adecuada para la comparación de rendimientos entre sistemas diferentes.

Bibliografía

(2004). Recognition of human actions. <http://www.nada.kth.se/cvap/actions/>. Accessed: 2015-11-15.

(2005). Actions as space-time shapes. <http://www.wisdom.weizmann.ac.il/~vision/SpaceTimeActions.html>. Accessed: 2016-01-05.

(2006). Inria xmas motion acquisition sequences. <http://4drepository.inrialpes.fr/public/viewgroup/6>. Accessed: 2015-10-08.

Aggarwal, J. and Ryoo, M. S. (2011). Human activity analysis: A review. *ACM Computing Surveys (CSUR)*, 43(3):16.

Bell, D. and Wang, H. (2000). A formalism for relevance and its application in feature subset selection. *Machine Learning*, 41(2):175–195.

Blum, A. L. and Langley, P. (1997). Selection of relevant features and examples in machine learning. *Artificial Intelligence*, 97(1–2):245 – 271.

Chen, C.-H., Liu, T.-L., Wang, Y.-S., Chu, H.-K., Tang, N. C., and Liao, H.-Y. M. (2015). Spatio-temporal learning of basketball offensive strategies. In *Proceedings of the 23rd Annual ACM Conference on Multimedia Conference*, pages 1123–1126. ACM.

Cilla, R., Patricio, M. A., Berlanga, A., and Molina, J. M. (2011). A probabilistic, discriminative and distributed system for the recognition of human actions from multiple views. *Neurocomputing*, 71(1):78–87.

Fish, B., Khan, A., Chehade, N., Chien, C., and Pottie, G. (2012). Feature selection based on mutual information for human activity recognition. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pages 1729–1732.

Gonen, M. and Alpaydm, E. (2011). Multiple kernel learning algorithms. *Journal of Machine Learning Research*, 12:2211–2268.

Gong, S. and Xiang, T. (2003). Recognition of group activities using dynamic probabilistic networks. In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pages 742–749. IEEE.

Gärdenfors, P. (1978). On the logic of relevance. *Synthese*, 37(3):351–367.

- Hakeem, A. and Shah, M. (2007). Learning, detection and representation of multi-agent events in videos. *Artificial Intelligence*, 171(8):586–605.
- Hongeng, S. and Nevatia, R. (2001). Multi-agent event recognition. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, volume 2, pages 84–91. IEEE.
- Intille, S. S. and Bobick, A. F. (2001). Recognizing planned, multiperson action. *Computer Vision and Image Understanding*, 81(3):414–445.
- John, G. H., Kohavi, R., Pfleger, K., et al. (1994). Irrelevant features and the subset selection problem. In *Machine Learning: Proceedings of the Eleventh International Conference*, pages 121–129.
- Khan, S. M. and Shah, M. (2005). Detecting group activities using rigidity of formation. In *Proceedings of the 13th annual ACM international conference on Multimedia*, pages 403–406. ACM.
- Koller, D. and Sahami, M. (1995). Toward optimal feature selection. In *In 13th International Conference on Machine Learning*, pages 284–292.
- Li, R. and Chellappa, R. (2010). Recognizing offensive strategies from football videos. In *Image Processing (ICIP), 2010 17th IEEE International Conference on*, pages 4585–4588. IEEE.
- Li, R., Chellappa, R., and Zhou, S. K. (2009). Learning multi-modal densities on discriminative temporal interaction manifold for group activity recognition. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 2450–2457. IEEE.
- Li, R., Chellappa, R., and Zhou, S. K. (2013a). Recognizing interactive group activities using temporal interaction matrices and their riemannian statistics. *International Journal of Computer Vision*, 101(2):305–328.
- Li, X., Hu, W., Shen, C., Zhang, Z., Dick, A., and Hengel, A. V. D. (2013b). A survey of appearance models in visual object tracking. *ACM Trans. Intell. Syst. Technol.*, 4(4):58:1–58:48.
- Liu, H. and Motoda, H. (2012). *Feature selection for knowledge discovery and data mining*, volume 454. Springer Science & Business Media.
- Liu, H., Sun, J., Liu, L., and Zhang, H. (2009). Feature selection with dynamic mutual information. *Pattern Recognition*, 42(7):1330 – 1339.
- Liu, X. and Chua, C.-S. (2006). Multi-agent activity recognition using observation decomposed hidden markov models. *Image and vision computing*, 24(2):166–175.
- Maldonado, S. and Weber, R. (2009). A wrapper method for feature selection using support vector machines. *Information Sciences*, 179(13):2208 – 2217. Special Section on High Order Fuzzy Sets.
- Mitchell, T. M. (1982). Generalization as search. *Artificial intelligence*, 18(2):203–226.

- Moeslund, T. B., Hilton, A., and Kruger, V. (2006). A survey of advances in vision-based human motion capture and analysis. *Computer vision and image understanding*, 104(2):90–126.
- Morariu, V. I. and Davis, L. S. (2011). Multi-agent event recognition in structured scenarios. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 3289–3296. IEEE.
- Narendra, P. M. and Fukunaga, K. (1977). A branch and bound algorithm for feature subset selection. *Computers, IEEE Transactions on*, C-26(9):917–922.
- Patricio, M. and Maravall, D. (2000). Segmentation of text and graphics/images using the gray-level histogram fourier transform. *Advances in Pattern Recognition*, pages 757–766.
- Pearl, J. (1984). *Heuristics addison-wesley. Reading, MA.*
- Peng, H., Long, F., and Ding, C. (2005). Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(8):1226–1238.
- Perše, M., Kristan, M., Kovačič, S., Vučkovič, G., and Perš, J. (2009). A trajectory-based analysis of coordinated team activity in a basketball game. *Computer Vision and Image Understanding*, 113(5):612–621.
- Poppe, R. (2010). A survey on vision-based human action recognition. *Image and vision computing*, 28(6):976–990.
- Pozo, A., García, J., and Patricio, M. (2012a). Inef12basketball dataset and the group behavior recognition issue. *Distributed Computing and Artificial Intelligence*, pages 151–160.
- Pozo, A., Gracia, J., Patricio, M. A., and Molina, J. M. (2011a). Group behavior recognition in context-aware systems. In *Advances in Computational Intelligence*, pages 645–652. Springer.
- Pozo, A., Gracia, J., Patricio, M. A., and Molina, J. M. (2011b). A structured representation to the group behavior recognition issue. In *User-Centric Technologies and Applications*, pages 47–57. Springer.
- Pozo, A., Patricio, M. A., García, J., and Molina, J. M. (2013). A feature selection approach to the group behavior recognition issue using static context information. *International Journal of Distributed Sensor Networks*, 2013.
- Pozo, A., Patricio Guisado, M. Á., García, J., Molina, J. M., and Refoyo, I. (2012b). Group behavior recognition issue, feature analysis on defending pick and roll basketball move.
- Siddique, B., Yacoob, Y., and Davis, L. (2009). Recognizing plays in american football videos. Technical report, Citeseer.
- Swears, E. and Hoogs, A. (2009). Learning and recognizing american football plays. *preservation*, 1(4):5–10.

- Vaswani, N., Roy-Chowdhury, A. K., and Chellappa, R. (2005). Shape activity: a continuous-state hmm for moving/deforming shapes with application to abnormal activity detection. *Image Processing, IEEE Transactions on*, 14(10):1603–1616.
- Vere, S. A. (1975). Induction of concepts in the predicate calculus. In *Proceedings of the Fourth International Joint Conference on Artificial Intelligence*, volume 1, pages 281–287.
- Viterbi, A. J. (1967). Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *Information Theory, IEEE Transactions on*, 13(2):260–269.
- Weinland, D., Ronfard, R., and Boyer, E. (2011). A survey of vision-based methods for action representation, segmentation and recognition. *Computer Vision and Image Understanding*, 115(2):224–241.
- Welch, L. R. (2003). Hidden markov models and the baum-welch algorithm. *IEEE Information Theory Society Newsletter*, 53(4):10–13.
- Winston, P. H. (1970). Learning structural descriptions from examples. Technical report, DTIC Document.
- Yang, H., Shao, L., Zheng, F., Wang, L., and Song, Z. (2011). Recent advances and trends in visual tracking: A review. *Neurocomputing*, 74(18):3823 – 3831.
- Yilmaz, A., Javed, O., and Shah, M. (2006). Object tracking: A survey. *Acm Computing Surveys (CSUR)*, 38(4):13.
- Zhang, D., Gatica-Perez, D., Bengio, S., and McCowan, I. (2006). Modeling individual and group actions in meetings with layered hmms. *Multimedia, IEEE Transactions on*, 8(3):509–520.
- Zhou, Y., Yan, S., and Huang, T. S. (2008). Pair-activity classification by bi-trajectories analysis. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE.